Faithful Simulation of Distributed Quantum Measurements With Applications in Distributed Rate-Distortion Theory

Touheed Anwar Atif, *Graduate Student Member, IEEE*, Mohsen Heidari, and S. Sandeep Pradhan, *Fellow, IEEE*

***Abstract*— We consider the task of faithfully simulating a distributed quantum measurement, wherein we provide a protocol for the three parties, Alice, Bob and Charlie, to simulate a repeated action of a distributed quantum measurement using a pair of non-product approximating measurements by Alice and Bob, followed by a stochastic mapping at Charlie. The objective of the protocol is to utilize minimum resources, in terms of classical bits needed by Alice and Bob to communicate their measurement outcomes to Charlie, and the common randomness shared among the three parties, while faithfully simulating independent repeated instances of the original measurement. To achieve this, we develop a mutual covering lemma and a technique for random binning of distributed quantum measurements, and, in turn, characterize a set of sufficient communication and common randomness rates required for asymptotic simulatability in terms of single-letter quantum information quantities. In the special case, where the Charlie's action is restricted to a deterministic mapping, we develop a one-shot performance characterization of the distributed faithful simulation problem. Furthermore, using these results we address a distributed quantum rate-distortion problem, where we characterize the achievable rate distortion region through a single-letter inner bound. Finally, via a technique of single-letterization of multi-letter quantum information quantities, we provide an outer bound for the rate-distortion region.**

***Index Terms*— Quantum measurement, distributed measurements, measurement compression, channel simulation, mutual covering, quantum Shannon theory.**

I. Introduction

MEASUREMENTS interface the intricate quantum world with the perceivable macroscopic classical world by associating a classical attribute to a quantum state. However, quantum phenomena, such as superposition, entanglement, and non-commutativity contribute to uncertainty in the measurement outcomes. A key concern, from an information-theoretic standpoint, is to quantify the amount of "relevant information" conveyed by a measurement about a quantum state.

Winter's measurement compression theorem [1] (also elaborated in [2]) quantifies the "relevant information" as the amount of resources needed to faithfully simulate the output of a quantum measurement applied on a given state in an asymptotic sense. Imagine that an agent (Alice) performs a measurement M on a quantum state ρ, and sends a set of classical bits to a receiver (Bob). Bob intends to *faithfully* recover the outcomes of Alice's measurements without having access to ρ, while preserving the correlation with the post-measured state of Alice's reference. One of the salient features of the measurement compression theorem is that it achieves the following asymptotic performance. If at least quantum mutual information ($I(X;R)$) amount of classical information and conditional entropy ($S(X|R)$) amount of common shared randomness are available, then one can achieve *faithful simulation* of the measurement M with respect to the quantum state ρ, where R denotes a reference of the quantum state, and X denotes the auxiliary register corresponding to the random measurement outcome. Wilde *et al.* [2] extended the measurement compression problem by considering additional resources available to each of the participating parties. One such formulation allows Bob to further process the information received from Alice using local private randomness. In analogy with [3], this problem formulation is referred to as non-feedback measurement simulation, while the former is termed as simulation with feedback. This quantified the benefit of private randomness in terms of enhancing the trade-off between classical bits communicated and common random bits consumed. In particular, the use of private randomness increases the requirement of classical communication bits, while reducing the common randomness constraint.

The problem of quantifying the information gain of a measurement has been studied extensively. Early works include [4]–[6]. Later on, Buscemi *et al.* [7]–[9] proposed the quantum mutual information with respect to a classical-quantum state as the measure to characterize the corresponding information gain. Subsequently, Berta *et al.* [10] provided a universal measurement compression theorem, generalizing the Winter's measurement compression theorem for arbitrary inputs. They identified the quantum mutual information of a measurement as the information gained by performing the

Manuscript received July 1, 2021; revised September 30, 2021; accepted October 18, 2021. Date of publication November 2, 2021; date of current version January 20, 2022. This work was supported by NSF under Grant CCF-1717523 and Grant CCF-2007878. An earlier version of this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT) in 2019 [DOI: 10.1109/ISIT.2019.8849691]. *(Corresponding author: S. Sandeep Pradhan.)*

Touheed Anwar Atif and S. Sandeep Pradhan are with the Department of Electrical Engineering and Computer Science (EECS), University of Michigan, Ann Arbor, MI 48109 USA (e-mail: touheed@umich.edu; pradhanv@umich.edu).

Mohsen Heidari is with the Department of Computer Science (CS), Purdue University, West Lafayette, IN 47907 USA (e-mail: mheidari@purdue.edu).

Communicated by M. M. Wilde, Associate Editor for Quantum Information Theory.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3124976.

Digital Object Identifier 10.1109/TIT.2021.3124976

measurement, independent of the input state on which it is performed. The proof was based on a new "classically coherent state merging protocol" - a variation of the quantum state merging protocol [11], [12], and the post-selection technique for quantum channels [13]. Recently, Anshu *et al.* [14] considered the problem of measurement compression with side information in the one-shot setting. They presented a protocol employing convex-split lemma for classical-quantum states [15], [16] and position based decoding [17], and bounded the communication in terms of smooth max and hypothesis testing relative entropies. On a similar note, Renes and Renner [18] studied the problem of sending classical messages in the presence of quantum side information in the one-shot setting. We direct an interested reader to [19], [20] for a detailed discussion and results pertaining to one-shot quantum information theory.

The measurement compression theorem [1] finds its applications in several quantum paradigms. It is a predecessor to the quantum reverse Shannon theorem [3], [21], [22], useful in determining the communication cost of the local purity distillation protocol [23]–[26], and also helpful in the first step of the so-called grandmother protocol [27] which involves distillation of entanglement from noisy bipartite states. This theorem was later used by Datta *et al.* [28] to develop a quantum-to-classical (q-c) rate-distortion theory. The problem involved lossy compression of a quantum information source into classical bits, with the task of compression performed by applying a measurement on the source. In this problem, the objective is to minimize the storage of the classical outputs resulting from the measurement, while being able to recover the quantum state (from classical bits) within a fixed level of distortion as measured by an observable. To achieve this, the authors in [29] advocated the use of the measurement compression protocol, and subsequently characterized the so-called rate-distortion function in terms of single-letter quantum mutual information quantities. The authors further established that by employing a naive approach of measuring individual output of the quantum source, and then applying Shannon's rate-distortion theory to compress the classical data obtained is insufficient to achieve optimal rates. Further, the problem of measurement compression in the presence of quantum side information was studied in [2]. The authors here combined the ideas from [1] and [30] to reduce the classical communication rate and common randomness needed to simulate a measurement in presence of quantum side information. Recently, authors in [14] came up with a completely different technique for analyzing the measurement simulation protocols, while considering the problem of quantum measurement compression with side information. They provide a protocol based on convex-split and position-based decoding, and bound rates from above in terms of smooth max and hypothesis testing relative entropies (defined in [14]).

In this work, we consider scenarios where the quantum measurements are performed in a distributed fashion on bipartite entangled states, and quantify "relevant information" for these distributed quantum measurements in an asymptotic sense. As shown in Fig. 1, a composite bipartite quantum system AB is made available to two agents, Alice and Bob, where they have access to the sub-systems A and B, respectively. Two separate measurements, one for each sub-system, are performed in a distributed fashion with no communication taking place between Alice and Bob. Imagine that there is a third party, Charlie, who is connected to Alice and Bob via two separate classical links. The objective of the three parties is to simulate the action of repeated independent measurements performed on many independent copies of the given composite state. To achieve this objective, Alice and Bob send classical bits to Charlie at rate R_1 and R_2, respectively. Further, pairwise common randomness at rates C_1 and C_2 are also shared between Alice and Charlie, and Bob and Charlie, respectively. Charlie performs classical processing of the received bits and common randomness. We study two settings, based on whether or not Charlie has access to private randomness. As an application of this quantification, we consider the quantum-to-classical distributed rate distortion problem where Charlie is allowed to use classical-to-quantum channels. In this work, we focus on memoryless quantum systems in finite-dimensional Hilbert spaces. We summarize the contributions of this work in the following:

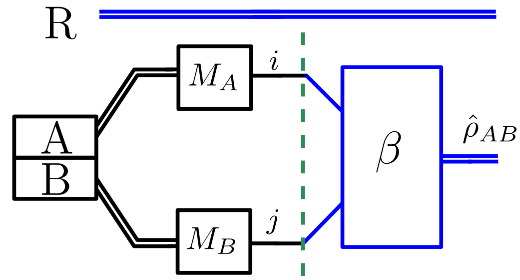

Fig. 1. The diagram of a distributed quantum measurement applied to a bipartite quantum system AB. A tensor product measurement $M_A \otimes M_B$ is performed on many copies of the observed quantum state. The outcomes of the measurements are given by two classical bit streams. The receiver functions as a classical-to-quantum channel β mapping the classical data to a quantum state.

- We formulate the problem of faithful simulation of distributed quantum measurements that can be decomposed as a convex-linear combination (incorporating Charlie's stochastic processing) of separable measurements, as stated in Definition 1. The asymptotic performance limit for this problem is given by the set of all communication rates (R_1, R_2) and all common randomness rates C_1 and C_2, referred to as the achievable rate region, under which the above-stated measurement is distributively simulated. We devise a distributed simulation protocol for this problem, and provide a quantum-information theoretic inner bound to the achievable rate region in terms of computable single-letter information quantities (see Theorem 2). This is the first main result of the paper.
- In the special case of the above problem formulation, where the Charlie's action is restricted to a deterministic mapping, we develop a one-shot performance characterization of the distributed faithful simulation problem (see Theorem 3). This characterization is based on a modular approach. As a corollary to this result, we develop a

characterization of an inner bound to the asymptotic performance limit (see Theorem 4).

- As an immediate application of our results on the simulation of distributed measurements, we develop an approach for a distributed quantum-to-classical rate distortion theory, where the objective is to reconstruct a quantum state at Charlie, with the quality of reconstruction measured using an additive distortion observable. The asymptotic performance limit is given by the set of all communication rate pairs (R_1, R_2) at which the distortion D is achieved. For the achievability part, we characterize an inner bound in terms of single-letter quantum mutual information quantities (see Theorem 5). This is the second main result of the paper. The classical version of this result is called the Berger-Tung inner bound [31].
- We then develop a technique for deriving converse bounds based on a combination of *tensor-product* and *direct-sum* Hilbert spaces (also referred to as a multi-particle system). Using this technique, we derive a single-letter outer-bound on the optimal rate distortion region (see Theorem 6), by converting a multi-letter expression into a single-letter expression. This is the third main result of the paper.

As was pointed out in [26], the measurement compression theorem [1] is a generalization of the classical reverse Shannon theorem [14] and can be viewed as a quantum-to-classical channel simulation problem. Similarly, the distributed measurement compression problem addressed in our work can be viewed as a distributed multi-party quantum-to-classical channel simulation problem. and can pave the way to considering the multi-party extensions of problems such as entanglement distillation and remote state preparation. Further, this work also develops new tools such as the mutual covering lemma and the mutual packing lemma which can be promising tools for many emerging quantum network applications. Moreover, in the recent applications of the distributed paradigms, a network of limited qubit-capacity quantum computers, connected through classical and quantum channels, are used to solve problems in a distributed manner by casting known centralized algorithms into their distributed versions [32]–[35].

The organization of the paper is as follows. In Section II, we set the notation and state requisite definitions. In Section III we state all the main results developed in this work. Toward developing the proof of these results, for pedagogical reasons, we first consider a special case in Section IV. For this special case, we restrict the processing at Charlie to a deterministic function and characterize the performance of a faithful simulation protocol in a one-shot setting. We achieve this by first obtaining a one-shot measurement compression theorem in a point-to-point setting (Theorem 7), wherein Bob is absent. Then we employ this result on the individual components (M_A and M_B) of the joint measurement M_{AB}, separately, to obtain a theorem characterizing the performance of a distributed measurement compression protocol (see Theorem 3). As a corollary, we further provide an asymptotic quantum information-theoretic inner bound to the achievable rate region of the distributed measurement compression problem (see Theorem 4). As a result, faithful simulation of M_A is possible when at least $nI(U; R_A)$ classical bits of communication and $nS(U|R_A)$ bits of common randomness are available between Alice and Charlie. Similarly, a faithful simulation of M_B is possible with $nI(V; R_B)$ classical bits of communication and $nS(V|R_B)$ bits of common randomness between Charlie and Bob, where R_A and R_B are purifications of the sub-systems A and B, respectively, and U and V denote the auxiliary registers corresponding to their measurement outcomes. The challenge here is that the direct use of single-POVM compression theorem for each individual POVMs, M_A and M_B, does not necessarily ensure a "distributed" faithful simulation of the overall measurement, M_{AB}. To accomplish this, we develop a Mutual Covering Lemma (see Lemma 4), which also helps in converting the information quantities in terms of the reference R of the joint state ρ_{AB}.

Further, an interesting aspect about the distributed setting is that one can further reduce the amount of classical communication by exploiting the statistical correlations between Alice's and Bob's measurement outcomes. The challenge here is that the classical outputs of the approximating POVMs (operating on n copies of the state ρ_{AB}) are not independent identically distributed (IID) sequences — rather they are codewords generated from random coding. For this we develop a proposition for mutual packing (Proposition 2), that characterizes the binning rates in term of single-letter information quantities. This issue also arises in classical distributed source coding problem which was addressed by Wyner-Ahlswede-Körner [31] by developing the Markov lemma and the Mutual packing lemma. The idea of binning in quantum setting has been explored from a different perspective in [30] and [36] for quantum data compression involving side information. Toward the end of the section, we also provide an example to illustrate the inner bound to the achievable rate region.

In Section V, we apply this special setting of the distributed measurement simulation with deterministic processing to the q-c distributed rate distortion problem. Since the proof of the inner bound of this rate distortion problem requires only the special case of distributed measurement simulation, this is another reason for providing the special case in the previous section.

In Section VI, we consider the non-feedback measurement compression problem for the point-to-point setting. The authors in [2] have discussed this formulation and provided a rate region with a proof of achievability and converse. However, in their proof, the authors assume two inequalities [2, Eq. 53 and 54], which may not necessarily be true [37] (further details are provided in Section VI). A stronger version of this theorem is also developed in [10] using a different technique, wherein the authors have extended the Winter's measurement compression for fixed independent and identically distributed inputs [1] to arbitrary inputs. Since the result is crucial for the distributed simulation problem with stochastic processing, to be proved in the next section (Section VII), we formally state the problem and provide an alternative proof of the direct part for completeness (see Theorem 8).

Finally, the above proof of non-feedback simulation in the point-to-point setting provides us with necessary tools for the next task, namely, distributed quantum measurement simulation with stochastic processing. The objective of incorporating the additional processing at the decoder is to reduce the required shared randomness. Our objective in the distributed problem, considered in Section III-B, was to simulate $M_A \otimes M_B$. We achieve this by proving that a pair of POVMs that can faithfully simulate M_A and M_B individually, can also faithfully simulate $M_A \otimes M_B$ (Lemma 4). However, it will be shown that, because of the presence of Charlie's stochastic processing, decoupling the current problem into two symmetric point-to-point problems is not feasible. Therefore, we perform a non-symmetric partitioning while being analytically tractable. Toward this we develop a non-product covering lemma (see Proposition 7). Moreover, we provide a single-letter achievable inner bound that is symmetric with respect to Alice and Bob. We conclude the paper with a few remarks in Section VIII.

II. Preliminaries

We here establish all our notations, briefly state few necessary definitions, and also provide Winter's theorem on measurement compression.

Notation: Given any natural number n, let the finite set $\{1, 2, \cdots, n\}$ be denoted by $[1, n]$. Let $\mathcal{B}(\mathcal{H})$ denote the algebra of all bounded linear operators acting on a finite-dimensional Hilbert space $\mathcal{H}$. Further, let $\mathcal{D}(\mathcal{H})$ denote the set of all unit trace positive operators acting on $\mathcal{H}$. Let I denote the identity operator. The trace distance between two operators A and B is defined as $\|A - B\|_1 \triangleq \operatorname{Tr}|A - B|$, where for any operator Λ we define $|\Lambda| \triangleq \sqrt{\Lambda^\dagger \Lambda}$. The von Neumann entropy of a density operator $\rho \in \mathcal{D}(\mathcal{H})$ is denoted by $S(\rho)$. The quantum mutual information for a bipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined as

$$I(A; B)_\rho \triangleq S(\rho_A) + S(\rho_B) - S(\rho_{AB}).$$

Given any ensemble $\{p_i, \rho_i\}_{i \in [1,m]}$, the Holevo information, as in [38], is defined as

$$\chi\big(\{p_i, \rho_i\}\big) \triangleq S\Big(\sum_i p_i \rho_i\Big) - \sum_i p_i S(\rho_i).$$

A positive operator-valued measure (POVM) acting on a Hilbert space $\mathcal{H}$ is a collection $M \triangleq \{\Lambda_x\}_{x \in \mathcal{X}}$ of positive operators in $\mathcal{B}(\mathcal{H})$ that form a resolution of the identity:

$$\Lambda_x \geqslant 0, \forall x \in \mathcal{X}, \qquad \sum_{x \in \mathcal{X}} \Lambda_x = I,$$

where $\mathcal{X}$ is a finite set. If instead of the equality above, the inequality $\sum_x \Lambda_x \leqslant I$ holds, then the collection is said to be a sub-POVM. A sub-POVM M can be completed to form a POVM, denoted by $[M]$, by adding the operator $\Lambda_0 \triangleq (I - \sum_x \Lambda_x)$ to the collection. Let Ψ^ρ_{RA} denote a purification of a density operator $\rho \in D(\mathcal{H}_A)$. Given a POVM $M \triangleq \{\Lambda^A_x\}_{x \in \mathcal{X}}$ acting on $\rho \in D(\mathcal{H}_A)$, the post-measurement state of the reference together with the classical outputs is represented by

$$(\mathrm{id} \otimes M)(\Psi^\rho_{RA}) \triangleq \sum_{x \in \mathcal{X}} |x\rangle\langle x| \otimes \operatorname{Tr}_A\{(I^R \otimes \Lambda^A_x)\Psi^\rho_{RA}\}. \quad (1)$$

Consider two POVMs $M_A = \{\Lambda^A_x\}_{x \in \mathcal{X}}$ and $M_B = \{\Lambda^B_y\}_{y \in \mathcal{Y}}$ acting on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Define $M_A \otimes M_B \triangleq \{\Lambda^A_x \otimes \Lambda^B_y\}_{x \in \mathcal{X}, y \in \mathcal{Y}}$. With this definition, $M_A \otimes M_B$ is a POVM acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. By $M^{\otimes n}$ denote the n-fold tensor product of the POVM M with itself.

Definition 1 (Joint Measurements): A POVM $M_{AB} \triangleq \{\Lambda^{AB}_z\}_{z \in \mathcal{Z}}$, acting on the joint state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, is said to have a separable decomposition with stochastic integration if there exist POVMs $M_A \triangleq \{\Lambda^A_u\}_{u \in \mathcal{U}}$ and $M_B \triangleq \{\Lambda^B_v\}_{v \in \mathcal{V}}$ and a stochastic mapping $P_{Z|U,V} : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$ such that

$$\Lambda^{AB}_z \triangleq \sum_{u,v} P_{Z|U,V}(z|u,v) \Lambda^A_u \otimes \Lambda^B_v, \quad \forall z \in \mathcal{Z},$$

where $\mathcal{U}, \mathcal{V}$ and $\mathcal{Z}$ are some finite sets. Further, if the mapping $P_{Z|U,V}$ is a deterministic function then the POVM is said to have a separable decomposition with deterministic integration.

Measurement Compression Theorem: Here, we provide a brief overview of the measurement compression theorem [1]. A key concern, from an information-theoretic standpoint, is to quantify the amount of "relevant information" conveyed by a measurement about a quantum state. Winter quantified "relevant information" by measuring the minimum amount of classical information bits needed to "simulate" the repeated action of a measurement M on a quantum state ρ. In this context, an agent (Alice) performs an approximating measurement $\tilde{M}^{(n)}$ on a quantum state $\rho^{\otimes n}$ and sends a set of classical bits to a receiver (Bob). In addition, Alice and Bob share some amount of common randomness. Bob intends to faithfully recover the outcomes of the original measurement M without having access to the quantum state based on the bits received from Alice and the common randomness. The objective is to minimize the rate of classical bits under the constraint that the approximating measurement $\tilde{M}^{(n)}$ is faithful to the actual measurement $M^{\otimes n}$ with respect to the state $\rho^{\otimes n}$. This is formally defined in the following.

Definition 2 (Faithful Simulation [2]): Given a sub-POVM $M \triangleq \{\Lambda_x\}_{x \in \mathcal{X}}$ acting on a Hilbert space $\mathcal{H}_A$ and a density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$, a sub-POVM $\tilde{M} \triangleq \{\tilde{\Lambda}_x\}_{x \in \mathcal{X}}$ acting on $\mathcal{H}_A$ is said to be ϵ-faithful to M with respect to ρ, for $\epsilon > 0$, if the following holds:

$$\Xi_\rho(M, \tilde{M}) \triangleq \sum_{x \in \mathcal{X}} \left\|\sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho}\right\|_1 + \operatorname{Tr}\left\{(I - \sum_x \Lambda_x)\rho\right\} + \operatorname{Tr}\left\{(I - \sum_x \tilde{\Lambda}_x)\rho\right\} \leqslant \epsilon. \quad (2)$$

Alternatively, one can complete the POVMs M and $\tilde{M}$ by associating $I - \sum_{x \in \mathcal{X}} \Lambda_x$ and $I - \sum_{x \in \mathcal{X}} \tilde{\Lambda}_x$ with additional symbols 0 and $\tilde{0}$, respectively, and thus obtaining POVMs $[M]$ and $[\tilde{M}]$, defined on $\mathcal{X} \bigcup \{0, \tilde{0}\}$. Stating the above definition for $[M]$ and $[\tilde{M}]$ gives the same as in [2, Definition 3]. Further, the above trace norm constraint can be equivalently expressed in terms of a purification of state ρ using the following lemma.

Lemma 1: [2, Lemma 4] For any state $\rho \in \mathcal{D}(\mathcal{H})$ with any purification Ψ^{ρ}_{RA}, and any pair of POVMs M and $\tilde{M}$ acting on $\mathcal{H}$, the following identity holds

$$\|(id\otimes M)(\Psi^{\rho}_{RA})-(id\otimes\tilde{M})(\Psi^{\rho}_{RA})\|_1=\sum_x\|\sqrt{\rho}(\Lambda_x-\tilde{\Lambda}_x)\sqrt{\rho}\|_1, \tag{3}$$

where Λ_x and $\tilde{\Lambda}_x$ are the operators associated with M and $\tilde{M}$, respectively.

Theorem 1: [1, Theorem 2] For any $\epsilon > 0$, any density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$, any POVM M acting on the Hilbert space $\mathcal{H}_A$, and for all sufficiently large n, there exists a collection of POVMs $\tilde{M}^{(n,\mu)}$ for $\mu \in [1, N]$, each acting on $\mathcal{H}_A^{\otimes n}$, and having at most 2^{nR} outcomes such that $\tilde{M}^{(n)} \triangleq \frac{1}{N}\sum_{\mu}\tilde{M}^{(n,\mu)}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ if

$$R > I(U;R)_\sigma, \quad \text{and} \quad \frac{1}{n}\log_2 N + R > S(U)_\sigma,$$

where $\sigma_{RU} \triangleq (id\otimes M)(\Psi^{\rho}_{RA})$.

Remark 1: A strong converse of the above result is also provided in [1, Theorem 8].

III. Main Results

In this section, we provide the main results of the paper.

A. Simulation of Distributed POVMs With Stochastic Processing

We begin by considering the simulation of distributed POVMs with stochastic processing. Consider a bipartite composite quantum system (A, B) represented by a Hilbert Space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let ρ_{AB} be a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$. Consider a joint measurement M_{AB} on the system. Imagine that three parties, named Alice, Bob and Charlie, are trying to collectively simulate the joint measurement, using two measurements, one applied on each sub-system. The resources available to these parties are: some amount of classical common randomness pairwise shared among them, and classical communication links of specified rates between Alice and Charlie, and Bob and Charlie. Alice and Bob perform measurements $\tilde{M}_A^{(n)} \triangleq \{\Lambda^A_{l_1}\}$ and $\tilde{M}_B^{(n)} \triangleq \{\Lambda^B_{l_2}\}$ on n copies of sub-systems A and B, respectively. The measurements are performed in a distributed fashion with no communication taking place between Alice and Bob. Based on their respective measurements and the common randomness, Alice and Bob send some classical bits to Charlie. Upon receiving these classical bits, Charlie applies a stochastic processing operation on them, given by $P(\cdot|l_1, l_2)$, and then wishes to produce an n-letter classical sequence. The objective is to construct n-letter measurements $\tilde{M}_A^{(n)}$ and $\tilde{M}_B^{(n)}$ that minimize the classical communication and common randomness bits while ensuring that the overall measurement induced by the action of the three parties is close to $M_{AB}^{\otimes n}$. Further, the operators of the given measurement M_{AB} admit a decomposition of the form given in Definition 1. We formally define the problem as follows.

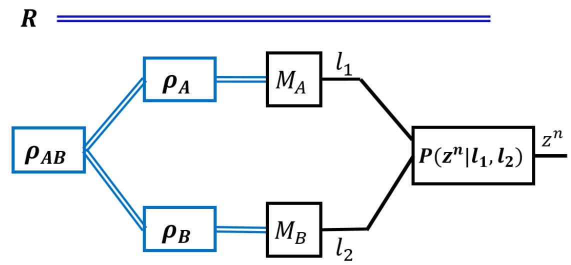

Fig. 2. The diagram depicting the distributed POVM simulation problem with stochastic processing. In this setting, Charlie additionally has access to unlimited private randomness.

1) Problem Formulation: The problem is defined in the following.

Definition 3 (Distributed Protocol): For a given finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, a distributed protocol with stochastic processing with parameters $(n, \Theta_1, \Theta_2, N_1, N_2)$ is characterized by

1) a collection of Alice's sub-POVMs $\tilde{M}_A^{(\mu_1)}, \mu_1 \in [1, N_1]$ each acting on $\mathcal{H}_A^{\otimes n}$ and with outcomes in $[1, \Theta_1]$.

2) a collection of Bob's sub-POVMs $\tilde{M}_B^{(\mu_2)}, \mu_2 \in [1, N_2]$ each acting on $\mathcal{H}_B^{\otimes n}$ and with outcomes in $[1, \Theta_2]$.

3) a collection of Charlie's classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ for all $l_1 \in [1, \Theta_1], l_2 \in [1, \Theta_2], z^n \in \mathcal{Z}^n$, $\mu_1 \in [1, N_1]$, and $\mu_2 \in [1, N_2]$.

The overall sub-POVM of this distributed protocol, given by $\tilde{M}_{AB}$, is characterized by the following operators:

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{l_1\in[1,\Theta_1],l_2\in[1,\Theta_2]} P^{(\mu_1,\mu_2)}(z^n|l_1,l_2) \times \Lambda^{A,(\mu_1)}_{l_1} \otimes \Lambda^{B,(\mu_2)}_{l_2}, \quad \forall z^n \in \mathcal{Z}^n,$$

where $\Lambda^{A,(\mu_1)}_{l_1}$ and $\Lambda^{B,(\mu_2)}_{l_2}$ are the operators corresponding to the sub-POVMs $\tilde{M}_A^{(\mu_1)}$ and $\tilde{M}_B^{(\mu_2)}$, respectively.

In the above definition, (Θ_1, Θ_2) determines the amount of classical bits communicated from Alice and Bob to Charlie, respectively. N_1 and N_2 denote the amount of pairwise common randomness. The classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ represent the action of Charlie on the received classical bits.

Definition 4 (Achievability): Given a POVM M_{AB} acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, and a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, a quadruple (R_1, R_2, C_1, C_2) is said to be achievable, if for all $\epsilon > 0$ and for all sufficiently large n, there exists a distributed protocol with stochastic processing with parameters $(n, \Theta_1, \Theta_2, N_1, N_2)$ such that its overall sub-POVM $\tilde{M}_{AB}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$ (see Definition 2), and

$$\frac{1}{n}\log_2 \Theta_i \leqslant R_i + \epsilon, \quad \text{and} \quad \frac{1}{n}\log_2 N_i \leqslant C_i + \epsilon, \quad i = 1, 2.$$

The set of all achievable quadruples (R_1, R_2, C_1, C_2) is called the achievable rate region.

2) Main Result: The following theorem provides an inner bound to the achievable rate region. The proof of the theorem is provided in Section VII, while some of the tools required for the proof are developed in Section VI.

Theorem 2: Given a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, and a POVM $M_{AB} = \{\Lambda_z^{AB}\}_{z\in\mathcal{Z}}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ having a separable decomposition with stochastic integration (as in Definition 1), a quadruple (R_1, R_2, C_1, C_2) is achievable if the following inequalities are satisfied:

$$\begin{aligned} R_1 &\geqslant I(U;RB)_{\sigma_1} - I(U;V)_{\sigma_3}, \\ R_2 &\geqslant I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \\ R_1+R_2 &\geqslant I(U;RB)_{\sigma_1} + I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \\ R_1+C_1 &\geqslant I(U;RZ)_{\sigma_3} - I(U;V)_{\sigma_3}, \\ R_2+C_2 &\geqslant I(V;RZ)_{\sigma_3} - I(U;V)_{\sigma_3}, \\ R_1+R_2+C_1 &\geqslant I(U;RZ)_{\sigma_3} + I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \\ R_1+R_2+C_2 &\geqslant I(V;RZ)_{\sigma_3} + I(U;RB)_{\sigma_1} - I(U;V)_{\sigma_3}, \\ R_1+R_2+C_1+C_2 &\geqslant I(UV;RZ)_{\sigma_3}, \end{aligned} \tag{4}$$

for some decomposition with POVMs $M_A = \{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B = \{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a stochastic map $P_{Z|U,V} : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$, where the above information quantities are computed for the auxiliary states $\sigma_1^{RUB} \triangleq (id_R \otimes M_A \otimes id_B)(\Psi_{RAB}^{\rho_{AB}}), \sigma_2^{RAV} \triangleq (id_R \otimes id_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, and $\sigma_3^{RUVZ} \triangleq \sum_{u,v,z} \sqrt{\rho_{AB}} \left(\Lambda_u^A \otimes \Lambda_v^B\right) \sqrt{\rho_{AB}} \otimes P_{Z|U,V}(z|u,v) \, |u\rangle\langle u| \otimes |v\rangle\langle v| \otimes |z\rangle\langle z|$, and $\Psi_{RAB}^{\rho_{AB}}$ is a purification[1] of ρ_{AB}.

Remark 2: An alternative characterization of the above rate region can be obtained in terms of Holevo information. For this, we use the canonical ensembles $\{\lambda_u^A, \hat{\rho}_u^A\}$, $\{\lambda_v^B, \hat{\rho}_v^B\}$ and $\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}$ defined as

$$\begin{aligned} &\lambda_u^A \triangleq \mathrm{Tr}\{\Lambda_u^A \rho_A\}, \quad \lambda_v^B \triangleq \mathrm{Tr}\{\Lambda_v^B \rho_B\}, \\ &\lambda_{uv}^{AB} \triangleq \mathrm{Tr}\{(\Lambda_u^A \otimes \Lambda_v^B)\rho_{AB}\}, \quad \text{and} \\ &\hat{\rho}_u^A \triangleq \frac{1}{\lambda_u^A}\sqrt{\rho_A}\Lambda_u^A\sqrt{\rho_A}, \quad \hat{\rho}_v^B \triangleq \frac{1}{\lambda_v^B}\sqrt{\rho_B}\Lambda_v^B\sqrt{\rho_B}, \\ &\hat{\rho}_{uv}^{AB} \triangleq \frac{1}{\lambda_{uv}^{AB}}\sqrt{\rho_{AB}}(\Lambda_u^A \otimes \Lambda_v^B)\sqrt{\rho_{AB}}. \end{aligned} \tag{5}$$

Note that the post-measurement states corresponding to the outcomes u and v are given by $(\hat{\rho}_u^A)^T, (\hat{\rho}_v^B)^T$ and $(\hat{\rho}_{uv}^{AB})^T$, where transposes are defined with respect to the eigenbasis of the corresponding density operators. This entails that the states $\hat{\rho}_u^A, \hat{\rho}_v^B$ and $\hat{\rho}_{uv}^{AB}$ defined above have the same spectrum as the states induced on the purifying reference R after the measurement. However, these canonical states are not on the same "operational level" as the latter. Further, we define the following ensemble $\{\lambda_z, \hat{\rho}_z\}$ as

$$\begin{aligned} \lambda_z &\triangleq \sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} \lambda_{uv}^{AB} P_{Z|UV}(z|u,v) \quad \text{and} \\ \hat{\rho}_z &\triangleq \sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} P_{UV|Z}(u,v|z)\hat{\rho}_{uv}^{AB}, \end{aligned}$$

with $P_{UV|Z}(u,v|z) = \lambda_{uv}^{AB} \cdot P_{Z|UV}(z|u,v)/\lambda_z$ for all $(u,v,z) \in \mathcal{U} \times \mathcal{V} \times \mathcal{Z}$. With this ensemble, we have $I(U;RB)_{\sigma_1} = \chi\left(\{\lambda_u^A, \hat{\rho}_u^A\}\right)$, $I(V;RA)_{\sigma_2} = \chi\left(\{\lambda_v^B, \hat{\rho}_v^B\}\right)$, and $I(UV;RZ)_{\sigma_3} = I(UV;Z) + \chi\left(\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}\right) - \chi\left(\{\lambda_z, \hat{\rho}_z\}\right)$.

[1]The information theoretic quantities remain independent of the purification used in their definitions.

B. One-Shot Simulation of Distributed POVMs With Deterministic Processing

We now consider simulation of distributed POVMs with deterministic processing. Recall from the discussion in Section I that the motivation behind the restriction to deterministic processing is that the proof becomes modular, and also forms a first pedagogical step towards the distributed simulation with stochastic processing (Theorem 2). Due to the modularity of the proof, we were able to develop a one-shot version of the proof. In the following, we state the problem formulation and provide the theorem statement.

1) Problem Formulation: In this formulation, Charlie's processing is restricted to a deterministic mapping. More precisely, in the $(n, \Theta_1, \Theta_2, N_1, N_2)$ protocol as defined in Definition 3, the Charlie's action is given by the collection of decoding maps $f^{(\mu_1,\mu_2)} : [1,\Theta_1] \times [1,\Theta_2] \to \mathcal{Z}^n$ for $\mu_1 \in [1,N_1], \mu_2 \in [1,N_2]$.

The overall sub-POVM of this distributed protocol, given by $\tilde{M}_{AB}$, is characterized by the following operators:

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{\substack{l_1\in[1,\Theta],\\ l_2\in[1,\Theta]}} \mathbb{1}_{\{f^{(\mu_1,\mu_2)}(l_1,l_2)=z^n\}}\Lambda_{l_1}^{A,(\mu_1)} \otimes \Lambda_{l_2}^{B,(\mu_2)} \tag{6}$$

$\forall z^n \in \mathcal{Z}^n$, where $\Lambda_{l_1}^{A,(\mu_1)}$ and $\Lambda_{l_2}^{B,(\mu_2)}$ are the operators corresponding to the sub-POVMs $\tilde{M}_A^{(\mu_1)}$ and $\tilde{M}_B^{(\mu_2)}$, respectively. The achievable rate region can also be defined in a correspondingly straightforward way.

2) Main Results: We now provide two theorems characterizing the performance of faithful simulation protocols, one in a one-shot and the other in an asymptotic quantum information theoretic settings which form our main results on faithful simulation of distributed measurements with deterministic processing. The proofs of these theorems are provided in Section IV-B and IV-C.

Theorem 3 (One-Shot Distributed Faithful Simulation): Consider a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a sub-POVM $M_{AB} \triangleq \{\Lambda_u^A \otimes \Lambda_v^B\}_{u\in\mathcal{U},v\in\mathcal{V}}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. Suppose there exists total subspace projectors Π_{ρ_A}, Π_{ρ_B}, and codeword subspace projectors $\{\Pi_u^A\}_{u\in\mathcal{U}}$, $\{\Pi_v^A\}_{v\in\mathcal{V}}$, acting on $\mathcal{H}_A$ and $\mathcal{H}_B, respectively$, satisfying:

$$\begin{aligned} &\mathrm{Tr}\{\Pi_{\rho_A}\hat{\rho}_u^A\} \geqslant 1-\epsilon_1, \quad \mathrm{Tr}\{\Pi_{\rho_B}\hat{\rho}_v^B\} \geqslant 1-\epsilon_2, \\ &\mathrm{Tr}\{\Pi_u^A\hat{\rho}_u^A\} \geqslant 1-\epsilon_1, \quad \mathrm{Tr}\{\Pi_v^B\hat{\rho}_v^B\} \geqslant 1-\epsilon_2, \end{aligned} \tag{7a}$$

$$\begin{aligned} &\mathrm{Tr}\{\Pi_{\rho_A}\} \leqslant D_1, \quad \mathrm{Tr}\{\Pi_{\rho_B}\} \leqslant D_2, \\ &\Pi_u^A\hat{\rho}_u^A\Pi_u^A \leqslant \frac{1}{d_1}\Pi_u^A, \quad \Pi_v^B\hat{\rho}_v^B\Pi_v^B \leqslant \frac{1}{d_2}\Pi_v^B, \end{aligned} \tag{7b}$$

$$\begin{aligned} &\Pi_{\rho_A}\rho_A\Pi_{\rho_A} \leqslant \rho_A, \quad \Pi_{\rho_B}\rho_B\Pi_{\rho_B} \leqslant \rho_B, \\ &\Pi_u^A\hat{\rho}_u^A\Pi_u^A \leqslant \hat{\rho}_u^A, \quad \Pi_v^B\hat{\rho}_v^B\Pi_v^B \leqslant \hat{\rho}_v^B, \end{aligned} \tag{7c}$$

$$\begin{aligned} &\Pi_{\rho_A}\rho_A\Pi_{\rho_A} \leqslant \frac{1}{F_1}\Pi_{\rho_A}, \quad \sqrt{\rho_A}^{-1}\Pi_{\rho_A}\sqrt{\rho_A}^{-1} \leqslant f_1\Pi_{\rho_A}, \\ &\Pi_{\rho_B}\rho_B\Pi_{\rho_B} \leqslant \frac{1}{F_2}\Pi_{\rho_B}, \quad \sqrt{\rho_B}^{-1}\Pi_{\rho_B}\sqrt{\rho_B}^{-1} \leqslant f_2\Pi_{\rho_B}, \end{aligned} \tag{7d}$$

where $\epsilon_i \in (0,\frac{1}{2})$, $0 < d_i < D_i$, and $f_i, F_i > 0$ for $i = 1,2$, and $\hat{\rho}_u^A$ and $\hat{\rho}_v^B$ are defined in (5). Let $\mathcal{W} \subseteq \mathcal{U} \times \mathcal{V}$ be an arbitrary set. Let K_1 and K_2 be arbitrary positive integers such that $|\mathcal{U}| \geqslant K_1$ and $|\mathcal{V}| \geqslant K_2$. Then for any $T_i \leqslant K_i$ for $i = 1, 2$, there exists a distributed protocol with deterministic processing for the finite set $\mathcal{U} \times \mathcal{V}$ and the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, with parameters $(1, T_1, T_2, N_1, N_2)$ such that

$$\Xi_{\rho^{AB}}(M_{AB}, \tilde{M}_{AB}) \leqslant \alpha_A + \alpha_B + \alpha_P,$$

where

$$\alpha_A(\epsilon_1, N_1, K_1) \triangleq \frac{2}{(1+\epsilon_1)\sqrt{N_1 K_1}} \sum_{u\in\mathcal{U}} \sqrt{\lambda_u^A} + \frac{2\epsilon_1}{\epsilon_1 + 1} + f(\epsilon_1, \theta_1) + 4D_1 N_1 \exp\left[-\frac{K_1 \epsilon_1^3 d_1 D_1^{-1}}{4\ln 2}\right] + 2\theta_1, \quad (8)$$

$$\alpha_B(\epsilon_2, N_2, K_2) \triangleq \frac{2}{(1+\epsilon_2)\sqrt{N_2 K_2}} \sum_{v\in\mathcal{V}} \sqrt{\lambda_v^B} + \frac{2\epsilon_2}{\epsilon_2 + 1} + f(\epsilon_2, \theta_2) + 4D_2 N_2 \exp\left[-\frac{K_2 \epsilon_2^3 d_2 D_2^{-1}}{4\ln 2}\right] + 2\theta_2, \quad (9)$$

$$\alpha_P(\epsilon_1, \epsilon_2, K_1, K_2, N_1, N_2, T_1, T_2, \mathcal{W}) \triangleq 2\alpha_A + 2\alpha_B + 2\lambda^{AB}(\mathcal{W}^c) + \frac{2}{(1+\epsilon_1)(1+\epsilon_2)}\left[\frac{\lambda_{\text{m}}^A \lambda_{\text{m}}^B |\mathcal{W}| K_1 K_2}{(1-\theta_1)(1-\theta_2)T_1 T_2} + \frac{K_1 W_A \lambda_{\text{m}}^A}{(1-\theta_1)T_1}\left(1 + \frac{\lambda_{\text{m}}^B K_2}{(1-\theta_2)}\right) + \frac{K_2 W_B \lambda_{\text{m}}^B}{(1-\theta_2)T_2} \times \left(1 + \frac{\lambda_{\text{m}}^A K_1}{(1-\theta_1)}\right)\right]\frac{f_1 f_2}{F_1 F_2}, \quad (10)$$

and $\lambda_{u,v}^{AB} \triangleq \text{Tr}(\rho^{AB}(\Lambda_u \otimes \Lambda_v))$ with marginals $(\lambda_u^A, \lambda_v^B)$, $W_A \triangleq \max_{v\in\mathcal{V}} |\{u : (u,v) \in \mathcal{W}\}|$, and $W_B \triangleq \max_{u\in\mathcal{U}} |\{v : (u,v) \in \mathcal{W}\}|$, $\lambda_{\text{m}}^A \triangleq \max_u \lambda_u^A$, and $\lambda_{\text{m}}^B \triangleq \max_v \lambda_v^B$, $\theta_1 \triangleq 1 - \sum_{u\in\mathcal{U}} \lambda_u^A$, $\theta_2 \triangleq 1 - \sum_{v\in\mathcal{V}} \lambda_v^B$, $f(\epsilon, \theta) \triangleq \left[4\sqrt{\epsilon} + 4\sqrt{\epsilon + 2\sqrt{\epsilon}} + 4\sqrt{2}(1-\theta)\sqrt{\epsilon + \sqrt{\epsilon}}\right]/(1+\epsilon)$, and $\lambda^{AB}(\mathcal{W}^c) \triangleq \sum_{(u,v)\in\mathcal{W}^c} \lambda_{u,v}^{AB}$.

Remark 3: Note that the terms α_A and α_B can be identified as the one-shot expressions for the errors induced in approximating each of the sub-POVMs $\{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $\{\Lambda_v^B\}_{v\in\mathcal{V}}$, using their respective approximations. This approximation employs the one-shot version of the measurement compression theorem (Theorem 7), which is developed as a part of the proof in Section IV-B. Within α_A, the exponential term corresponds to the error probability that the approximating operators do not constitute a valid sub-POVMs in random coding, the term involving square-root of the probabilities corresponds to the classical soft covering error, and the term $f(\epsilon, \theta)$ corresponds to the error incurred because of the use of gentle measurement lemma with regard to the total subspace and codeword subspace projectors. Likewise, the term α_P captures the additional error introduced by compressing the classical outcomes of the above distributed measurement using the technique of binning. The binning is used to reduce the rate of transmission by exploiting the classical correlations present in the measurement outcomes, using a many-to-one transformation. The information lost in this transformation is recovered at the receiver using a relation modeled by a bipartite sub-graph $\mathcal{W}$ of $\mathcal{U} \times \mathcal{V}$. The twice of $\alpha_A + \alpha_B$ within α_P captures the effect of binning on the event corresponding to not being able to cover the sources using the approximating sub-POVMs. $\lambda^{AB}(\mathcal{W}^c)$ captures the event where under the original sub-POVM, the measurement outcomes do not satisfy the above set relation. The final term captures the error due to binning of the approximating sub-POVMs.

As a corollary to the above theorem, we obtain the following asymptotic inner bound to the achievable rate region.

Theorem 4: Given a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a POVM $M_{AB} \triangleq \{\Lambda_z^{AB}\}_{z\in\mathcal{Z}}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, and having a separable decomposition with deterministic integration (as in Definition 1), a quadruple (R_1, R_2, C_1, C_2) is achievable if the following inequalities are satisfied:

$$R_1 \geqslant I(U; RB)_{\sigma_1} - I(U; V)_{\sigma_3}, \quad (11a)$$

$$R_2 \geqslant I(V; RA)_{\sigma_2} - I(U; V)_{\sigma_3}, \quad (11b)$$

$$R_1 + R_2 \geqslant I(U; RB)_{\sigma_1} + I(V; RA)_{\sigma_2} - I(U; V)_{\sigma_3}, \quad (11c)$$

$$R_1 + C_1 \geqslant S(U|V)_{\sigma_3}, \quad (11d)$$

$$R_2 + C_2 \geqslant S(V|U)_{\sigma_3}, \quad (11e)$$

$$R_1 + R_2 + C_1 \geqslant I(V; RA)_{\sigma_2} + S(U|V)_{\sigma_3}, \quad (11f)$$

$$R_1 + R_2 + C_2 \geqslant I(U; RB)_{\sigma_1} + S(V|U)_{\sigma_3}, \quad (11g)$$

$$R_1 + R_2 + C_1 + C_2 \geqslant S(U, V)_{\sigma_3}, \quad (11h)$$

for some decomposition with POVMs $M_A = \{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B = \{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a function $g : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$, where the information quantities are computed for the auxiliary states $\sigma_1^{RUB} \triangleq (id_R \otimes M_A \otimes id_B)(\Psi_{RAB}^{\rho_{AB}})$, $\sigma_2^{RAV} \triangleq (id_R \otimes id_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, and $\sigma_3^{RUV} \triangleq (id_R \otimes M_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, with $\Psi_{RAB}^{\rho_{AB}}$ being a purification of ρ_{AB}.

Remark 4: An alternative characterization of the above rate region can be obtained in terms of Holevo information. Using the canonical ensemble, we obtain

$$I(U; RB)_{\sigma_1} = S(RB)_{\sigma_1} - S(RB|U)_{\sigma_1} = S(\sum_{u\in\mathcal{U}} \lambda_u^A \hat{\rho}_u^A) - \sum_{u\in\mathcal{U}} \lambda_u^A S(\hat{\rho}_u^A) = \chi\left(\{\lambda_u^A, \hat{\rho}_u^A\}\right),$$

where the second equality follows by noting $S(RB)_{\sigma_1} = S(\rho_A)$, $\rho_A = \sum_{u\in\mathcal{U}} \lambda_u^A \hat{\rho}_u^A$, and using the result from [39, Eq. 11.54]. Similarly, we get $I(V; RA)_{\sigma_2} = \chi\left(\{\lambda_v^B, \hat{\rho}_v^B\}\right)$. Also, $I(U; V)_{\sigma_3}$, and $S(U, V)_{\sigma_3}$ are equal to the classical mutual information and joint entropy with respect to the joint distribution $\{\lambda_{uv}^{AB}\}$, respectively.

C. Distributed Rate-Distortion Theory

As an application of faithful simulation of distributed measurements (Theorem 4), we consider the distributed extension of q-c rate distortion coding [28]. This problem is a quantum counterpart of the classical distributed source coding. In this setting, consider a memoryless bipartite quantum source, characterized by $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Alice and Bob have access to sub-systems A and B, characterized by $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and $\rho_B \in \mathcal{D}(\mathcal{H}_B)$, respectively, where $\rho_A = \text{Tr}_B\{\rho_{AB}\}$ and $\rho_B = \text{Tr}_A\{\rho_{AB}\}$. They both perform a measurement on n copies of their sub-systems and send the classical bits to Charlie. Upon receiving the classical bits sent by Alice and Bob, a reconstruction state is produced by Charlie. The objective of Charlie is to produce a reconstruction of the source ρ_{AB} within a targeted distortion threshold which is measured by a given distortion observable.

1) Problem Formulation: We first formulate this problem as follows. For any quantum information source, characterized by $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, denote its purification by $\Psi^{\rho_{AB}}_{RAB}$.

Definition 5 (q-c Source Coding Setup): A q-c source coding setup is characterized by a triple $(\Psi^{\rho_{AB}}_{RAB}, \mathcal{H}_{\hat{X}}, \Delta)$, where $\Psi^{\rho_{AB}}_{RAB} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$ is a purification of ρ_{AB}, $\mathcal{H}_{\hat{X}}$ is a reconstruction Hilbert space, and $\Delta \in \mathcal{B}(\mathcal{H}_R \otimes \mathcal{H}_{\hat{X}})$, which satisfies $\Delta \geqslant 0$, is a distortion observable.

Next, we formulate the action of Alice, Bob and Charlie by the following definition.

Definition 6 (q-c Protocol): An (n, Θ_1, Θ_2) q-c protocol for a given input and reconstruction Hilbert spaces $(\mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{H}_{\hat{X}})$ is defined by POVMs $M_A^{(n)}$ and $M_B^{(n)}$ acting on $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$ with Θ_1 and Θ_2 number of outcomes, respectively, and a set of reconstruction states $S_{i,j} \in \mathcal{D}(\mathcal{H}_{\hat{X}}^{\otimes n})$ for all $i \in [1, \Theta_1], j \in [1, \Theta_2]$.

The overall action of Alice, Bob and Charlie, as a q-c protocol, on a quantum source ρ_{AB} is given by the following operation

$$\mathcal{N}_{A^nB^n \mapsto \hat{X}^n} : \rho_{AB}^{\otimes n} \mapsto \sum_{i,j} \mathrm{Tr}\{(\Lambda_i^A \otimes \Lambda_j^B)\rho_{AB}^{\otimes n}\} \; S_{i,j}, \quad (12)$$

where $\{\Lambda_i^A\}$ and $\{\Lambda_j^B\}$ are the operators of the POVMs $M_A^{(n)}$ and $M_B^{(n)}$, respectively. With this notation and given a q-c source coding setup as in Definition 5, the distortion of a $(n = 1, \Theta_1, \Theta_2)$ q-c protocol is measured as

$$d(\rho_{AB}, \mathcal{N}_{AB\mapsto\hat{X}}) \triangleq \mathrm{Tr}\left\{\Delta\left((\mathrm{id}_R \otimes \mathcal{N}_{AB\mapsto\hat{X}})(\Psi^{\rho_{AB}}_{RAB})\right)\right\}.$$

For an n-letter protocol, we use symbol-wise average distortion observable defined as

$$\Delta^{(n)} = \frac{1}{n}\sum_{i=1}^{n} \Delta_{R_i\hat{X}_i} \otimes I_{R\hat{X}}^{\otimes[n]\setminus i}, \quad (13)$$

where $\Delta_{R_i\hat{X}_i}$ is understood as the observable Δ acting on the ith instance space $\mathcal{H}_{R_i} \otimes \mathcal{H}_{\hat{X}_i}$ of the n-letter space $\mathcal{H}_R^{\otimes n} \otimes \mathcal{H}_{\hat{X}}^{\otimes n}$. With this notation, the distortion for an (n, Θ_1, Θ_2) q-c protocol is given by

$$d(\rho_{AB}^{\otimes n}, \mathcal{N}_{A^nB^n\mapsto\hat{X}^n}) \triangleq \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id} \otimes \mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi^{\rho_{AB}}_{R^nA^nB^n})\right\},$$

where $\Psi^{\rho_{AB}}_{R^nA^nB^n}$ is the n-fold tensor product of $\Psi^{\rho_{AB}}_{RAB}$ which is the given purification of the source.

The authors in [28] studied the point-to-point setup of the above formulation wherein Bob is absent. They considered a special distortion observable of the form $\Delta = \sum_{\hat{x}\in\hat{\mathcal{X}}} \Delta_{\hat{x}} \otimes |\hat{x}\rangle\langle\hat{x}|$, where $\Delta_{\hat{x}} \geqslant 0$ acts on the reference Hilbert space and $\hat{\mathcal{X}}$ is the reconstruction alphabet (please see [28, Sec. 4] for more details). In this paper, we allow Δ to be any non-negative and bounded operator acting on the appropriate Hilbert spaces. Moreover, we allow for the use of any c-q reconstruction mapping as the action of Charlie.

Definition 7 (Achievability): For a q-c source coding setup $(\Psi^{\rho_{AB}}_{RAB}, \mathcal{H}_{\hat{X}}, \Delta)$, a rate-distortion triplet (R_1, R_2, D) is said to be achievable, if for all $\epsilon > 0$ and all sufficiently large n, there exists an (n, Θ_1, Θ_2) q-c protocol satisfying

$$\frac{1}{n}\log_2 \Theta_i \leqslant R_i + \epsilon, \quad i = 1, 2,$$
$$d(\rho_{AB}^{\otimes n}, \mathcal{N}_{A^nB^n\mapsto\hat{X}^n}) \leqslant D + \epsilon,$$

where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is defined as in (12). The set of all achievable rate-distortion triplets (R_1, R_2, D) is called the achievable rate-distortion region.

Our objective is to characterize the achievable rate-distortion region using single-letter information quantities.

2) Main Result: An Inner Bound: We provide an inner bound to the achievable rate-distortion region which is stated in the following theorem. We employ a q-c protocol based on a randomized faithful simulation strategy involving a time sharing classical random variable Q that is independent of the quantum source. This can be viewed as a conditional version of the faithful simulation problem considered in Section III-B. The proof of the theorem in provided in Section V.

Theorem 5: For a q-c source coding setup $(\Psi^{\rho_{AB}}_{RAB}, \mathcal{H}_{\hat{X}}, \Delta)$, any rate-distortion triplet (R_1, R_2, D) satisfying the following inequalities is achievable

$$\begin{aligned}
R_1 &\geqslant I(U;RB|Q)_{\sigma_1} - I(U;V|Q)_{\sigma_3},\\
R_2 &\geqslant I(V;RA|Q)_{\sigma_2} - I(U;V|Q)_{\sigma_3},\\
R_1 + R_2 &\geqslant I(U;RB|Q)_{\sigma_1} + I(V;RA|Q)_{\sigma_2} - I(U;V|Q)_{\sigma_3},\\
D &\geqslant d(\rho_{AB}, \mathcal{N}_{AB\mapsto\hat{X}}),
\end{aligned}$$

for POVM of the form $M_{AB} = \sum_{q\in\mathcal{Q}} P_Q(q) M_A^q \otimes M_B^q$, where for every $q \in \mathcal{Q}$, $M_A^q \triangleq \{\Lambda_u^{A,q}\}_{u\in\mathcal{U}}$ and $M_B^q \triangleq \{\Lambda_v^{B,q}\}_{v\in\mathcal{V}}$ are POVMs acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, and reconstruction states $\{S_{u,v,q}\}$ with each state in $\mathcal{D}(\mathcal{H}_{\hat{X}})$, and some finite sets $\mathcal{U}, \mathcal{V}$ and $\mathcal{Q}$. The quantum mutual information quantities are computed according to the auxiliary states $\sigma_1^{RUBQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(id_R \otimes M_A^q \otimes id_B)(\Psi^{\rho_{AB}}_{RAB}) \otimes |q\rangle\langle q|$, $\sigma_2^{RAVQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(id_R \otimes id_A \otimes M_B^q)(\Psi^{\rho_{AB}}_{RAB}) \otimes |q\rangle\langle q|$, and $\sigma_3^{RUVQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(id_R \otimes M_A^q \otimes M_B^q)(\Psi^{\rho_{AB}}_{RAB}) \otimes |q\rangle\langle q|$, where (U, V) represents the output of M_{AB}, and $\mathcal{N}_{AB\mapsto\hat{X}} : \rho_{AB} \mapsto \sum_{u,v,q} P_Q(q)\,\mathrm{Tr}\{(\Lambda_u^{A,q} \otimes \Lambda_v^{B,q})\rho_{AB}\}\; S_{u,v,q}$.

Remark 5: Note that for the auxiliary state σ_1, we have

$$\begin{aligned}
\sigma_1^{RQ} &= \mathrm{Tr}_{UB}\{\sigma_1^{RUBQ}\}\\
&= \sum_q P_Q(q)\,\mathrm{Tr}_{UAB}\left\{\sum_{u\in\mathcal{U}}\{(I_{RB} \otimes \Lambda_u^q)(\Psi^{\rho_{AB}}_{RAB})\} \otimes |u\rangle\langle u|\right\}\\
&\quad \otimes |q\rangle\langle q|\\
&= \sum_q P_Q(q)\,\mathrm{Tr}_{AB}\left\{\left\{(I_{RB} \otimes \sum_{u\in\mathcal{U}} \Lambda_u^q)(\Psi^{\rho_{AB}}_{RAB})\right\}\right\} \otimes |q\rangle\langle q|\\
&= \rho_R \otimes \sum_q P_Q(q)\,|q\rangle\langle q|,
\end{aligned}$$

which gives $I(R;Q)_{\sigma_1} = 0$. Similar statements hold for the states σ_2 and σ_3.

One can observe that the rate region in Theorem 5 matches in form with the classical Berger-Tung region when ρ_{AB} is a mixed state of a collection of orthogonal pure states. Note that the rate region is an inner bound for the set of all achievable rates. The single-letter characterization of the set of achievable rates is still an open problem even in the classical setting. Some progress has been made recently on this problem which provides an improvement over Berger-Tung rate region [40].

3) Main Result: An Outer Bound: In this section, we provide an outer bound for the achievable rate-distortion region. The proof of this theorem is provided in Section V.

Theorem 6: Given a q-c source coding setup $(\Psi^{\rho_{AB}}_{RAB}, \mathcal{H}_{\hat{X}}, \Delta)$, if any triplet (R_1, R_2, D) is achievable, then the following inequalities must be satisfied

$$R_1 \geqslant I(W_1; R|W_2, Q)_\sigma, \tag{14a}$$

$$R_2 \geqslant I(W_2; R|W_1, Q)_\sigma, \tag{14b}$$

$$R_1 + R_2 \geqslant I(W_1, W_2; R|Q)_\sigma, \tag{14c}$$

$$D \geqslant \mathrm{Tr}\left\{\Delta\sigma^{R\hat{X}}\right\}, \tag{14d}$$

for some state $\sigma^{W_1W_2RQ\hat{X}}$ which can be written as

$$\sigma^{W_1W_2QR\hat{X}} = (\mathrm{id} \otimes \mathcal{N}_{AB\mapsto W_1W_2Q\hat{X}})(\Psi^{\rho_{AB}}_{RAB}),$$

where W_1, W_2 and Q represent auxiliary quantum states, and $\mathcal{N}_{AB\mapsto W_1W_2Q\hat{X}}$ is a quantum test channel with $I(R;Q)_\sigma = 0$.

Remark 6: One may question the computability of the outer bound provided in Theorem 6. The computability of this bound depends on the dimensionality of the auxiliary space $\mathcal{H}_Q$ defined in the theorem. Currently, we are unable to bound the dimension of the Hilbert space $\mathcal{H}_Q$, but aim to provide one in our future work. As a matter of fact, the current outer bounds for the equivalent classical distributed rate distortion problem still suffers from the computability issue. The first outer bound to the classical problem was provided in [31] and a recent substantial improvement was made by authors in [41]. Both of these bounds suffer from the absence of cardinality bounds on at least one of the variables used, and hence cannot be claimed to be computable using finite resources.

IV. Proofs: Distributed Simulation of POVMs With Deterministic Processing

A. Overview of Proof Technique and an Illustrative Example

Before providing a proof in the next section, we briefly discuss two corner points of the rate region with respect to the common randomness available. To reduce the number of free parameters, let $C \triangleq C_1 + C_2$. Firstly, consider the regime where the sum rate $(R_1 + R_2)$ is at its minimum achievable, i.e., equation (11c) is active. This requires the largest amount of common randomness, given by the constraint $C \geqslant S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$. Next, let us consider the regime where $C = 0$. This implies $R_1 + R_2 \geqslant S(U, V)_{\sigma_3}$. This regime corresponds to the quantum measurement $M_A \otimes M_B$ followed by classical Slepian-Wolf compression [42]. Fig. 3 demonstrates the achievable rate region in these cases.

We encounter two challenges in developing the single-letter inner bound to the achievable rate region as stated in Theorem 4: 1) The direct use of single-POVM compression theorem, proved using random coding arguments as in [1], for each individual POVMs, M_A and M_B, does not necessarily ensure a "distributed" faithful simulation for the overall measurement, $M_A \otimes M_B$. This issue is unique to the quantum settings. One of the contributions of this work is to prove this when the two sources A and B are not necessarily independent, i.e., $\rho_{AB} \neq \rho_A \otimes \rho_B$ (see Lemma 4).

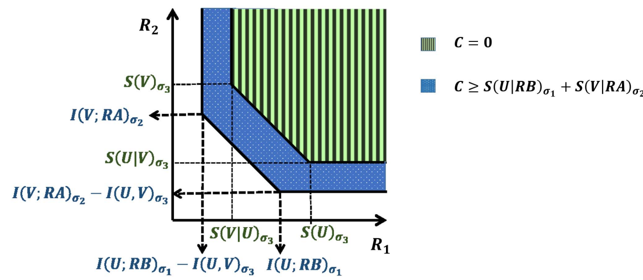

Fig. 3. The inner bound to the achievable rate region given in Theorem 4 at two planes: 1) with no common randomness, i.e., $C = 0$ (green color), and 2) with at least $S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$ amount of common randomness (blue color). As a result, the latter region contains the former.

2) The classical outputs of the approximating POVMs (operating on n copies of the source) are not *independently and identically distributed* (IID) sequences - rather they are codewords generated from random coding. The Slepian-Wolf scheme [42] (also referred to as *binning* in the literature) is developed for distributed compression of IID source sequences. Applicability of such an approach to the problem requires that the classical outputs produced from the two approximating POVMs are jointly typical with high probability. This issue also arises in classical distributed source coding problem which was addressed by Wyner-Ahlswede-Korner by developing the Markov Lemma and the Mutual Packing Lemma (Lemma 12.1 and 12.2 in [43]). Building upon these ideas, we develop quantum-classical counterparts of these lemmas for the multi-user quantum measurement simulation problem (see the discussion in Section VII-A.2 and Proposition 2).

Let us consider an example to illustrate the above inner bound.

Example 1: Suppose the composite state ρ_{AB} is described using one of the Bell states on $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$\rho^{AB} \triangleq \frac{1}{2}(|00\rangle_{AB} + |11\rangle_{AB})(\langle 00|_{AB} + \langle 11|_{AB}).$$

Since $\pi^A = \mathrm{Tr}_B\, \rho^{AB}$ and $\pi^B = \mathrm{Tr}_A\, \rho^{AB}$, Alice and Bob would perceive each of their particles in maximally mixed states $\pi^A = \frac{I^A}{2}$ and $\pi^B = \frac{I^B}{2}$, respectively. Upon receiving the quantum state, the two parties wish to independently measure their states, using identical POVMs M_A and M_B, given by $\left\{\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -|\right\}$. Alice and Bob together with Charlie are trying to simulate the action of $M_A \otimes M_B$ using the classical communication and common randomness as the resources available to them (as described earlier). We compute the constraints given in Theorem 4.

Considering the first constraint from (11a), we evaluate σ_1^{UB} as

$$\sigma_1^{UB} = \frac{1}{4}\big(|0\rangle\langle 0|_U \otimes |0\rangle\langle 0|_B + |1\rangle\langle 1|_U \otimes |1\rangle\langle 1|_B + |2\rangle\langle 2|_U \otimes |+\rangle\langle +|_B + |3\rangle\langle 3|_U \otimes |-\rangle\langle -|_B \big),$$

where the vectors $\{|0\rangle_U, |1\rangle_U, |2\rangle_U, |3\rangle_U\}$ denote a set of orthogonal states on the space $\mathcal{H}_U$. Based on this state, we get

$$S(\sigma_1^{RUB}) = S(\sigma_1^{UB}) = 2, \ \ S(\sigma_1^{RB}) = S(\sigma_1^{B}) = 1, \quad S(\sigma_1^{U}) = 2.$$

This gives $I(U;RB)_{\sigma_1}$ to be equal to 1 bit. Similarly, from the symmetry of the example, we also get $I(V;RA)_{\sigma_2}$ to be equal to 1 bit. Similarly, we can evaluate σ_3^{UV} as

$$\sigma_3^{UV} = \left(\frac{1}{8}\sum_{i=0}^{3} |i\rangle\langle i|_U \otimes |i\rangle\langle i|_V + \frac{1}{16}\sum_{i=0}^{3}\sum_{j=i+2}^{i+4} |i\rangle\langle i|_U \otimes |\text{mod }(j,4)\rangle\langle \text{mod }(j,4)|_V \right),$$

which gives

$$S(U,V)_{\sigma_3} = 3.5 \quad \text{and} \quad I(U;V)_{\sigma_3} = 0.5.$$

Therefore, we can write the constraints given in Theorem 4 as

$$R_1 \geqslant 0.5, \quad R_2 \geqslant 0.5, \quad R_1 + R_2 \geqslant 1.5, \quad R_1 + C_1 \geqslant 1.5, \quad R_2 + C_2 \geqslant 1.5, \quad R_1 + R_2 + C_1 \geqslant 2.5,$$
$$R_1 + R_2 + C_2 \geqslant 2.5, \quad \text{and} \quad R_1 + R_2 + C_1 + C_2 \geqslant 3.5.$$

Consider the case when $C = C_1 + C_2 \geqslant 2$ is available. By approximating M_A and M_B individually, we receive a gain of 1 bit, decreasing the rate from $S(U)_{\sigma_1} = 2$ bits to $I(U;RB)_{\sigma_1} = 1$ bit and similarly from $S(V)_{\sigma_2} = 2$ bits to $I(V;RA)_{\sigma_2} = 1$ bit. Binning of these approximating POVMs (as discussed in Section (VII-A.2)), gives an additional gain of half a bit, which is characterized by $I(U;V)_{\sigma_3} = 0.5$, thus giving us the achievable sum-rate of 1.5 bits.

B. Proof of Theorem 3

We begin the proof of the theorem by restating the measurement compression theorem (Theorem 1) in a one-shot quantum information theoretic setting. This restatement allows us to develop a one-shot mutual covering lemma, which is a crucial part of the current proof. The theorem is stated as follows:

Theorem 7 (One-Shot Point-to-Point Faithful Simulation): Consider a density operator $\rho \in \mathcal{D}(\mathcal{H})$ and a sub-POVM $M \triangleq \{\Lambda_x\}_{x\in\mathcal{X}}$ acting on $\mathcal{H}$, and let $\{\lambda_x, \hat{\rho}_x\}_{x\in\mathcal{X}}$ be the canonical ensemble[2] of M with respect to ρ. Suppose there exists a total subspace projector Π_ρ and codeword subspace projectors $\{\Pi_x\}_{x\in\mathcal{X}}$ acting on $\mathcal{H}$ satisfying:

$$\text{Tr}\{\Pi_\rho \hat{\rho}_x\} \geqslant 1 - \epsilon \tag{15a}$$
$$\text{Tr}\{\Pi_x \hat{\rho}_x\} \geqslant 1 - \epsilon \tag{15b}$$
$$\text{Tr}\{\Pi_\rho\} \leqslant D \tag{15c}$$
$$\Pi_x \hat{\rho}_x \Pi_x \leqslant \frac{1}{d}\Pi_x \tag{15d}$$
$$\Pi_x \hat{\rho}_x \Pi_x \leqslant \hat{\rho}_x \tag{15e}$$
$$\Pi_\rho \rho \Pi_\rho \leqslant \rho, \tag{15f}$$

where $\epsilon \in (0, \frac{1}{2})$, $0 < d < D$. Then there exists a collection of sub-POVMs $\tilde{M}^{(\mu)}$ for $\mu \in [1, N]$ each with at most K outcomes, with $K \leqslant |\mathcal{X}|$, and acting on $\mathcal{H}$ such that

$$\Xi_\rho(M, \tilde{M}) \leqslant \frac{2}{(1+\epsilon)\sqrt{NK}} \sum_{x\in\mathcal{X}} \sqrt{\lambda_x} + \frac{2\epsilon}{\epsilon + 1} + f(\epsilon, \theta) + 4DN \exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2} \right] + 2\theta,$$

where $\tilde{M} \triangleq \frac{1}{N}\sum_\mu \tilde{M}^{(\mu)}$, $\theta \triangleq 1 - \sum_{x\in\mathcal{X}} \lambda_x$, and $f(\epsilon,\theta) \triangleq \left[4\sqrt{\epsilon} + 4\sqrt{\epsilon + 2\sqrt{\epsilon}} + 4\sqrt{2}(1-\theta)\sqrt{\epsilon + \sqrt{\epsilon}} \right]/(1-\epsilon)$. □

Proof: The proof is provided in Appendix A. □

Moving ahead with the proof of the current theorem, assume that the operators of the original sub-POVM $M_{AB} = M_A \otimes M_B$ are denoted by $\{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $\{\Lambda_v^B\}_{v\in\mathcal{V}}$, respectively, where $\mathcal{U}$ and $\mathcal{V}$ are two finite sets. The proof follows by constructing a protocol for faithful simulation of $M_A \otimes M_B$. We start by generating the canonical ensembles[3] corresponding to M_A and M_B. Let Π_{ρ_A} and Π_{ρ_B} denote the total projectors for marginal density operators ρ_A and ρ_B, respectively. Also, for any $u \in \mathcal{U}$ and $v \in \mathcal{V}$, let Π_u^A and Π_v^B denote the codeword projectors. Let the canonical ensembles be $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$. For each $u \in \mathcal{U}$ and $v \in \mathcal{V}$ define

$$\tilde{\rho}_u^{A'} \triangleq \Pi_{\rho_A}\Pi_u^A \hat{\rho}_u^A \Pi_u^A \Pi_{\rho_A}, \quad \tilde{\rho}_v^{B'} \triangleq \Pi_{\rho_B}\Pi_v^B \hat{\rho}_v^B \Pi_v^B \Pi_{\rho_B}. \tag{16}$$

With the notation above, define $\sigma^{A'}$ and $\sigma^{B'}$as

$$\sigma^{A'} \triangleq \frac{1}{(1-\theta_1)} \sum_{u\in\mathcal{U}} \lambda_u^A \tilde{\rho}_u^{A'}, \quad \sigma^{B'} \triangleq \frac{1}{(1-\theta_2)} \sum_{v\in\mathcal{V}} \lambda_v^B \tilde{\rho}_v^{B'}. \tag{17}$$

Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigenstates of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than ϵ_1/D_A and ϵ_2/D_B, respectively. Lastly, define

$$\tilde{\rho}_u^A \triangleq \hat{\Pi}^A \tilde{\rho}_u^{A'} \hat{\Pi}^A, \quad \text{and} \quad \tilde{\rho}_v^B \triangleq \hat{\Pi}^B \tilde{\rho}_v^{B'} \hat{\Pi}^B, \tag{18}$$

for all $u \in \mathcal{U}$, and $v \in \mathcal{V}$ and $\sigma^A = \hat{\Pi}^A \sigma^{A'} \hat{\Pi}^A$, $\sigma^B = \hat{\Pi}^B \sigma^{B'} \hat{\Pi}^B$.

1) Construction of Random POVMs: In what follows, we construct two random POVMs one for each encoder. Fix positive integers K_1, K_2, N_1 and N_2. Let $\mu_1 \in [1, N_1]$ denote the common randomness shared between the first encoder and the decoder, and let $\mu_2 \in [1, N_2]$ denote the common randomness shared between the second encoder and the decoder. For each $\mu_1 \in [1, N_1]$ and $\mu_2 \in [1, N_2]$, randomly and independently select $K_1 \times K_2$ pairs denoted by $(U^{(\mu_1)}(l), V^{(\mu_2)}(k))$ from the set $\mathcal{U} \times \mathcal{V}$ according to the distribution:

$$\mathbb{P}\left((U^{(\mu_1)}(l), V^{(\mu_2)}(k)) = (u,v) \right) = \frac{\lambda_u^A \lambda_v^B}{(1-\theta_1)(1-\theta_2)}, \tag{19}$$

[2]Note that $\{\lambda_x\}_{x\in\mathcal{X}}$ is a sub-probability vector, i.e., a vector of non-negative real numbers whose sum is not greater than 1.

[3]Note that $\{\lambda_u^A\}_{u\in\mathcal{U}}$ and $\{\lambda_v^B\}_{v\in\mathcal{V}}$ are sub-probability vectors.

for $u \in \mathcal{U}, v \in \mathcal{V}$. Let $\mathcal{C}^{(\mu_1,\mu_2)}$ denote the collection $\{U^{(\mu_1)}(l), V^{(\mu_2)}(k)\}_{\{l\in[1,K_1],k\in[1,K_2]\}}$. Construct operators[4]

$$A_u^{(\mu_1)} \triangleq \gamma_u^{(\mu_1)}\left(\sqrt{\rho_A}^{-1}\tilde{\rho}_u^A\sqrt{\rho_A}^{-1}\right) \quad \text{and}$$
$$B_v^{(\mu_2)} \triangleq \zeta_v^{(\mu_2)}\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_v^B\sqrt{\rho_B}^{-1}\right), \tag{20}$$

where

$$\gamma_u^{(\mu_1)} \triangleq \frac{(1-\theta_1)}{(1+\epsilon_1)K_1}|\{l : U^{(\mu_1)}(l) = u\}| \quad \text{and}$$
$$\zeta_v^{(\mu_2)} \triangleq \frac{(1-\theta_2)}{(1+\epsilon_2)K_2}|\{k : V^{(\mu_2)}(k) = v\}|. \tag{21}$$

Let $\mathbb{1}_{\{\text{sP-1}\}}$ denote the indicator random variable corresponding to the event that $\{A_u^{(\mu_1)} : u \in \mathcal{U}\}$ forms a sub-POVM for all $\mu_1 \in [1, N_1]$. Similarly define $\mathbb{1}_{\{\text{sP-2}\}}$ with regard to $\{B_v^{(\mu_2)} : v \in \mathcal{V}\}$. If $\mathbb{1}_{\{\text{sP-1}\}} = 1$, then, for each $\mu_i \in [1, N_i]$ construct $M_i^{(\mu_i)}$, for $i = 1, 2$, as in the following:

$$M_1^{(\mu_1)} \triangleq \{A_u^{(\mu_1)} : u \in \mathcal{U}\}, \quad M_2^{(\mu_2)} \triangleq \{B_v^{(\mu_2)} : v \in \mathcal{V}\}.$$

These collections $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ are completed using the operators $A_{0_U}^{(\mu_1)} \triangleq I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}$ and $B_{0_V}^{(\mu_2)} \triangleq I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}$, and these operators are associated with symbols 0_U and 0_V. In the case of the complementary event, i.e., $\mathbb{1}_{\{\text{sP-i}\}} = 0$, we define $M_i^{(\mu_i)} \triangleq \{I\}$, for $i = 1, 2$, and denote the output as 0_U or 0_V, respectively. Hence by construction $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ are sub-POVMs for all $\mu_i \in [1, N_i]$, for $i = 1, 2$. For a fixed $\{\mathcal{C}^{(\mu_1,\mu_2)}\}_{\mu_1\in[1,N_1],\mu_2\in[1,N_2]}$, the probability distribution P induced on $(\mathcal{U} \cup \{0_U\}) \times (\mathcal{V} \cup \{0_V\})$ has the following salient features.

$$P\{(u, v)\} = \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},$$

if $(u, v) \in \mathcal{U} \times \mathcal{V}$, and

$$\begin{aligned} &P\big((\mathcal{U} \cup \{0_U\}) \times (\mathcal{V} \cup \{0_V\})\backslash(\mathcal{U} \times \mathcal{V})\big) \\ &= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\left(1 - \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u,v}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\right) \\ &\quad + \left(1 - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right), \end{aligned}$$

where $\Omega_{u,v}$ is defined as

$$\Omega_{u,v} \triangleq \text{Tr}\left\{\sqrt{\rho_A \otimes \rho_B}^{-1}(\tilde{\rho}_u^A \otimes \tilde{\rho}_v^B)\sqrt{\rho_A \otimes \rho_B}^{-1}\rho_{AB}\right\}. \tag{22}$$

Binning of POVMs: We introduce the quantum counterpart of the so-called *binning* technique which has been widely used in the context of classical distributed source coding. Fix positive integers (T_1, T_2) and choose a (μ_1, μ_2) pair. For each symbol $u \in \mathcal{U}$ assign an index from $[1, T_1]$ randomly and uniformly, such that the assignments for different sequences are done independently. Perform a similar random and independent assignment for all $v \in \mathcal{V}$ with indices chosen from $[1, T_2]$. Repeat this assignment for every $\mu_1 \in [1, N_1]$ and $\mu_2 \in [1, N_2]$. For each $i \in [1, T_1]$ and $j \in [1, T_2]$, let $\mathcal{B}_1^{(\mu_1)}(i)$ and $\mathcal{B}_2^{(\mu_2)}(j)$ denote the i^{th} and the j^{th} bins, respectively. More precisely, $\mathcal{B}_1^{(\mu_1)}(i)$ is the set of all u symbols with assigned index equal to i, and similar is $\mathcal{B}_2^{(\mu_2)}(j)$. Define the following operators:

$$\Gamma_i^{A,(\mu_1)} \triangleq \sum_{u\in\mathcal{B}_1^{(\mu_1)}(i)} A_u^{(\mu_1)}, \quad \Gamma_j^{B,(\mu_2)} \triangleq \sum_{v\in\mathcal{B}_2^{(\mu_2)}(j)} B_v^{(\mu_2)},$$

for all $i \in [1, T_1]$ and $j \in [1, T_2]$. Using these operators, we form the following collection:

$$M_A^{(\mu_1)} \triangleq \{\Gamma_i^{A,(\mu_1)}\}_{i\in[1,T_1]}, \quad M_B^{(\mu_2)} \triangleq \{\Gamma_j^{B,(\mu_2)}\}_{j\in[1,T_2]}. \tag{23}$$

Note that if $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ are sub-POVMs, then so are $M_A^{(\mu_1)}$ and $M_B^{(\mu_2)}$. This is due to the relations

$$\sum_i \Gamma_i^{A,(\mu_1)} = \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}, \quad \text{and} \quad \sum_j \Gamma_j^{B,(\mu_2)} = \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}.$$

To make $M_A^{(\mu_1)}$ and $M_B^{(\mu_2)}$ complete, we define $\Gamma_0^{A,(\mu_1)}$ and $\Gamma_0^{B,(\mu_2)}$ as $\Gamma_0^{A,(\mu_1)} = I - \sum_i \Gamma_i^{A,(\mu_1)}$ and $\Gamma_0^{B,(\mu_2)} = I - \sum_j \Gamma_j^{B,(\mu_2)}$, respectively.[5] Now, we intend to use the completions $[M_A^{(n,\mu_1)}]$ and $[M_B^{(n,\mu_2)}]$ as the POVMs for each encoder. In event that $\mathbb{1}_{\text{sP-i}} = 0$, for $i = 1, 2$, then the symbols 0_U and 0_V are mapped to 0. Also, note that the effect of the binning is in reducing the communication rates from $(\log(K_1 + 1), \log(K_2 + 1))$ to $(\log(T_1 + 1), \log(T_2 + 1))$.

Decoder Mapping: Note that the operators $\{A_u^{(\mu_1)} \otimes B_v^{(\mu_2)}\}_{u\in\mathcal{U},v\in\mathcal{V}}$ are used to simulate $M_A \otimes M_B$. Binning can be viewed as partitioning of the set of classical outcomes into bins. Suppose an outcome (U, V) occurred in the measurement process. Then, if the bins are small enough, one might be able to recover the outcomes by knowing the bin numbers. For that we create a decoder that takes as an input a pair of bin numbers and produces a pair of symbols (U, V). More precisely, we define a mapping $F^{(\mu_1,\mu_2)}$, for (μ_1, μ_2), acting on the outputs of $[M_A^{(\mu_1)}]\otimes[M_B^{(\mu_2)}]$ as follows. On observing (μ_1, μ_2) and the classical indices $(i, j) \in [1, T_1] \times [1, T_2]$ communicated by the encoders, the decoder populates

$$D_{i,j}^{(\mu_1,\mu_2)} \triangleq \Big\{(u, v) \in \mathcal{C}^{(\mu_1,\mu_2)} : (u, v) \in \mathcal{W} \text{ and } (u, v) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)\Big\},$$

where $\mathcal{W}$ is an arbitrary subset of $\mathcal{U}\times\mathcal{V}$. For every $\mu_l \in [1, N_l]$, for $l = 1, 2$, and $i \in [1, K_1]$ and $j \in [1, K_2]$, define the function $F^{(\mu_1,\mu_2)}(i, j) = (u, v)$ if (u, v) is the only element of $D_{i,j}^{(\mu_1,\mu_2)}$; otherwise $F^{(\mu_1,\mu_2)}(i, j) = (0_U, 0_V)$. Further, $F^{(\mu_1,\mu_2)}(i, j) = (0_U, 0_V)$ for $i = 0$ or $j = 0$. With this mapping, we form the following collection of operators, denoted by $\tilde{M}_{AB}$,

$$\begin{aligned}\tilde{\Lambda}_{u,v}^{AB} \triangleq &\frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1N_2}\sum_{\mu_1=1}^{N_1}\sum_{\mu_2=1}^{N_2}\sum_{(i,j):F^{(\mu_1,\mu_2)}(i,j)=(u,v)}\Gamma_i^{A,(\mu_1)}\otimes\Gamma_j^{B,(\mu_2)} \\ &+ (1 - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})(I \otimes I)\mathbb{1}_{\{(u,v)=(0_U,0_V)\}},\end{aligned}$$

[4]The inverse used in $\sqrt{\rho}^{-1}$ refers to the generalized inverse as defined in [38, Section 5.6].

[5]Note that $\Gamma_0^{A,(\mu_1)} = I - \sum_i \Gamma_i^{A,(\mu_1)} = I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}$ and $\Gamma_0^{B,(\mu_2)} = I - \sum_j \Gamma_j^{B,(\mu_2)} = I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}$.

$\forall (u,v) \in (\mathcal{U} \bigcup \{0_U\}) \times (\mathcal{V} \bigcup \{0_V\})$. Note that by construction $\tilde{M}_{AB}$ is a sub-POVM.

2) Analysis of POVM and Trace Distance: We show that $\tilde{M}_{AB}$ is a sub-POVM that is faithful to the sub-POVM $M_A \otimes M_B$, with respect to ρ_{AB}. More precisely, we provide a bound on

$$G_{\rho_{AB}} \triangleq \Xi(M_{AB}, \tilde{M}_{AB}). \tag{24}$$

Step 1 ($M_1^{\mu_1}$ and $M_2^{(\mu_2)}$ Are Sub-POVMs and Individually Approximating): As a first step, one can show that $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ individually approximate the corresponding POVMs in the expected sense. More precisely the following lemma holds.

Lemma 2: For the POVM ensemble described above, we have

$$\mathbb{E}(\Xi_{\rho_A}(M_A, M_1)) \leqslant \alpha_A(\epsilon_1, K_1, N_1),$$
$$\mathbb{E}(\Xi_{\rho_B}(M_B, M_2)) \leqslant \alpha_B(\epsilon_2, K_2, N_2),$$

where $M_1 \triangleq \frac{1}{N_1}\sum_{\mu_1} M_1^{(\mu_1)}$, and $M_2 \triangleq \frac{1}{N_2}\sum_{\mu_2} M_2^{(\mu_2)}$.

Proof: Follows from the proof of Theorem 7, as the assumptions of that theorem (which M_A and M_B have to satisfy) are met as a part of the current theorem statement (see (7a-7c)). □

Step 2 (Isolating the Effect of Un-Binned Approximating Measurements: In this step, we separate out the effect of un-binned approximating measurements from G in (24). This is done by adding and subtracting an appropriate term within the trace norm and applying triangle inequality, which bounds G as $G \leqslant S_1 + S_2$, where

$$S_1 \triangleq \Bigg\| (\mathrm{id} \otimes [M_A] \otimes [M_B])(\Psi^{\rho}_{RAB}) - \frac{1}{N_1 N_2}\sum_{\mu_1,\mu_2} (\mathrm{id} \otimes [M_1^{(\mu_1)}] \otimes [M_2^{(\mu_2)}])(\Psi^{\rho}_{RAB}) \Bigg\|_1,$$
$$S_2 \triangleq \Bigg\| \frac{1}{N_1 N_2}\sum_{\mu_1,\mu_2} (\mathrm{id} \otimes [M_1^{(\mu_1)}] \otimes [M_2^{(\mu_2)}])(\Psi^{\rho}_{RAB}) - (\mathrm{id} \otimes [\tilde{M}_{AB}])(\Psi^{\rho}_{RAB}) \Bigg\|_1, \tag{25}$$

where S_1 captures the effect of using approximating sub-POVMs M_1 and M_2 instead of the actual sub-POVMs M_A and M_B, while S_2 captures the error introduced by binning these approximating sub-POVMs. Before we proceed further, we provide the following lemma which will be useful in the rest of the paper.

Lemma 3: Given a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$, a sub-POVM $M_Y \triangleq \{\Lambda_y^B : y \in \mathcal{Y}\}$ acting on $\mathcal{H}_B$, for some set $\mathcal{Y}$, and any Hermitian operator Γ^A acting on $\mathcal{H}_A$, we have

$$\sum_{y\in\mathcal{Y}} \left\| \sqrt{\rho_{AB}}\left(\Gamma^A \otimes \Lambda_y^B\right)\sqrt{\rho_{AB}} \right\|_1 \leqslant \left\| \sqrt{\rho_A}\,\Gamma^A \sqrt{\rho_A} \right\|_1, \tag{26}$$

with equality if $\sum_{y\in\mathcal{Y}} \Lambda_y^B = I$, where $\rho_A \triangleq \mathrm{Tr}_B\{\rho_{AB}\}$.

Proof: The proof is provided in Appendix B-A. □

Next, we provide a bound on S_1 using the following Mutual Covering Lemma.

Lemma 4: (Mutual Covering Lemma) Suppose a sub-POVM $\hat{M}_X$ is ϵ_X-faithful to M_X with respect to ρ_X, and a sub-POVM $\hat{M}_Y$ is ϵ_Y-faithful to M_Y with respect to ρ_Y, where $\rho_X = \mathrm{Tr}_Y\{\rho_{XY}\}$ and $\rho_Y = \mathrm{Tr}_X\{\rho_{XY}\}$. Then the sub-POVM $\hat{M}_X \otimes \hat{M}_Y$ is $(\epsilon_X + \epsilon_Y)$-faithful to the POVM $M_X \otimes M_Y$ with respect to ρ_{XY}.

Proof: The proof is provided in the Appendix B-B. □

Using Lemma 4 with $\rho_{XY} = \rho_{AB}$, $\hat{M}_X = \frac{1}{N_1}\sum_{\mu_1} M_1^{(\mu_1)}$, $\hat{M}_Y = \frac{1}{N_2}\sum_{\mu_2} M_2^{(\mu_2)}$, $M_X = [M_A]$ and $M_Y = [M_B]$, and Lemma 2, we have $\mathbb{E}(S_1) \leqslant \alpha_A(\epsilon_1, N_1, K_1) + \alpha_B(\epsilon_2, N_2, K_2)$. For later convenience, we state the following lemma which will be used in analyzing the binning operation:

Lemma 5: We have

$$\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} \Bigg| \lambda^{AB}_{u,v} - \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2} \gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\mathbb{1}_{\{\mathrm{sP}-1\}}\mathbb{1}_{\{\mathrm{sP}-2\}} \Bigg| + \mathbb{1}_{\{\mathrm{sP}-1\}}\mathbb{1}_{\{\mathrm{sP}-2\}}\Big(1 - \frac{1}{N_1N_2}\sum_{u,v}\sum_{\mu_1,\mu_2}\gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\Big) + (1 - \mathbb{1}_{\{\mathrm{sP}-1\}}\mathbb{1}_{\{\mathrm{sP}-2\}}) \leqslant S_1, \tag{27}$$

where $\Omega_{u,v}$ is defined as in (22).

Proof: The proof follows from Lemma 2 in [2]. □

Step 3 (Analyzing the Effect of Binning): In this step, we provide an upper bound on S_2. For $(u,v) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)$, define $e^{(\mu_1,\mu_2)}(u,v) \triangleq F^{(\mu_1,\mu_2)}(i,j)$. For any $(u,v) \notin \mathcal{C}^{(\mu_1,\mu_2)}$ define $e^{(\mu_1,\mu_2)}(u,v) = (0_U, 0_V)$. Note that $e^{(\mu_1,\mu_2)}$ captures the overall effect of the binning followed by the decoding function $F^{(\mu_1,\mu_2)}$. For all $u \in \mathcal{U}$ and $v \in \mathcal{V}$, let $\Phi_{u,v} \triangleq |u,v\rangle\langle u,v|$. With this notation, we simplify S_2 using the following proposition.

Proposition 1: S_2 can be simplified as

$$S_2 = S_3 + \frac{2}{N_1}\sum_{\mu_1}\mathrm{Tr}\left(\left(I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}\right)\rho_A\right) + \frac{2}{N_2}\sum_{\mu_2}\mathrm{Tr}\left(\left(I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}\right)\rho_B\right) + 2\left(2 - \mathbb{1}_{\{\text{sP-1}\}} - \mathbb{1}_{\{\text{sP-2}\}}\right),$$

where

$$S_3 \triangleq \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} \left\|\Phi_{u,v} - \Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1 \gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}.$$

Proof: The proof is provided in Appendix C-A. □

In the next proposition we provide a bound on the expectation of S_2.

Proposition 2 (Mutual Packing): We have

$$\mathbb{E}[S_2] \leqslant \alpha_P(\epsilon_1, \epsilon_2, K_1, K_2, N_1, N_2, T_1, T_2, \mathcal{W}).$$

Proof: The proof is provided in Appendix C-B. □

Combining the results from the mutual covering and mutual packing lemmas we obtain

$$G \leqslant \alpha_A + \alpha_B + \alpha_P.$$

This completes the proof of the theorem.

C. Proof of Theorem 4

We develop a proof as a corollary to Theorem 3. Assume that the operators of the original POVM M_{AB} are decomposed as

$$\Lambda_z^{AB} \triangleq \sum_{u,v} \mathbb{1}_{\{g(u,v)=z\}} \Lambda_u^A \otimes \Lambda_v^B, \ \forall z \in \mathcal{Z}, \tag{28}$$

for some POVMs M_A and M_B with operators denoted by $\{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $\{\Lambda_v^B\}_{v\in\mathcal{V}}$, respectively, and for some function $g: \mathcal{U}\times\mathcal{V} \to \mathcal{Z}$ where $\mathcal{U}, \mathcal{V}$ and $\mathcal{Z}$ are three finite sets. In what follows, we show the existence of an (n, T_1, T_2, N_1, N_2) distributed protocol with the associated sub-POVM $\tilde{M}_{AB}^{(n)}$ that is ϵ-faithful to M_{AB} with respect to $\rho_{AB}^{\otimes n}$ (according to Definition 2), where $\epsilon > 0$ can be made arbitrarily small for all sufficiently large n. More precisely, we plan to show that

$$\sum_{z^n} \Big\| \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{z^n}^{AB} - \tilde{\Lambda}_{z^n}^{AB}\right) \sqrt{\rho_{AB}^{\otimes n}} \Big\|_1 \leqslant \epsilon. \tag{29}$$

Next we claim that it is sufficient to show that there exists a distributed protocol for the finite set $\mathcal{U}\times\mathcal{V}$ and the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, with parameters (n, T_1, T_2, N_1, N_2) such that the associated sub-POVM $\tilde{M}_{AB}^{(n)} = \{\tilde{\Lambda}_{u^n,v^n}^{AB}\}_{u^n\in\mathcal{V}^n, v^n\in\mathcal{V}^n}$ satisfies $\Xi_{\rho_{AB}^{\otimes n}}(M_A^{\otimes n} \otimes M_B^{\otimes n}, \tilde{M}_{AB}^{(n)}) \leqslant \epsilon$. This is because one can always apply the function $g(\cdot,\cdot)$ componentwise on (u^n, v^n) to yield a sub-POVM with operators

$$\tilde{\Lambda}_{z^n} \triangleq \sum_{u^n\in\mathcal{U}^n} \sum_{v^n\in\mathcal{V}^n} \mathbb{1}_{\{g^n(u^n,v^n)=z^n\}} \tilde{\Lambda}_{u^n,v^n}^{AB}, \quad \forall z^n \in \mathcal{Z}^n,$$

that satisfies the constraint (29) as

$$\begin{aligned}
&\sum_{z^n} \Bigg\| \sum_{u^n,v^n} \mathbb{1}_{\{g^n(u^n,v^n)=z^n\}} \Bigg(\sqrt{\rho_{AB}^{\otimes n}} (\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \\
&\qquad\qquad - \tilde{\Lambda}_{u^n,v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \Bigg) \Bigg\|_1 \\
&\leqslant \sum_{z^n} \sum_{u^n,v^n} \mathbb{1}_{\{g^n(u^n,v^n)=z^n\}} \Bigg\| \sqrt{\rho_{AB}^{\otimes n}} (\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \\
&\qquad\qquad - \tilde{\Lambda}_{u^n,v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \Bigg\|_1 \\
&= \sum_{u^n,v^n} \Bigg\| \sqrt{\rho_{AB}^{\otimes n}} (\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \tilde{\Lambda}_{u^n,v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \Bigg\|_1 .
\end{aligned}$$

Fix three free parameters $\delta > 0$, $\epsilon_1 > 0$, and $\epsilon_2 > 0$. We make the following identification with regard to Theorem 3. (a) Let $\rho_{AB} \leftrightarrow \rho_{AB}^{\otimes n}$, $M_A \leftrightarrow M_A^{\otimes n}$, and $M_B \leftrightarrow M_B^{\otimes n}$, which implies that $\lambda_u^A \leftrightarrow \lambda_{u^n}^A$, $\lambda_v^B \leftrightarrow \lambda_{v^n}^B$ and $\lambda_{uv}^{AB} \leftrightarrow \lambda_{u^nv^n}^{AB}$. (b) Let $\mathcal{U} \leftrightarrow \mathcal{T}_\delta^{(n)}(U)$, $\mathcal{V} \leftrightarrow \mathcal{T}_\delta^{(n)}(V)$, and $\mathcal{W} \leftrightarrow \mathcal{T}_\delta^{(n)}(U,V)$, where $\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{T}_\delta^{(n)}(V)$ and $\mathcal{T}_\delta^{(n)}(U,V)$ are the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively. (c) Furthermore, let $\Pi_{\rho_A} \leftrightarrow \Pi_{\rho_A,\delta}$ $\Pi_{\rho_B} \leftrightarrow \Pi_{\rho_B,\delta}$, $\Pi_u^A \leftrightarrow \Pi_{u^n,\delta}^A$, and $\Pi_v^B \leftrightarrow \Pi_{v^n,\delta}^B$, where $\Pi_{\rho_A,\delta}$ and $\Pi_{\rho_B,\delta}$ denote the δ-typical projectors (as in [39, Def. 15.1.3]) for marginal density operators ρ_A and ρ_B, respectively.[6] Also, for any $u^n \in \mathcal{T}_\delta^{(n)}(U)$ and $v^n \in \mathcal{T}_\delta^{(n)}(V)$, let $\Pi_{u^n,\delta}^A$ and $\Pi_{v^n,\delta}^B$ denote the strong conditional typical projectors (as in [39, Def. 15.2.4]) for the canonical ensembles $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$, respectively.

With the above identification, and using the property of typical sets and typical projectors, we now find the values of the variables $D_1, D_2, d_1, d_2, F_1, F_2, f_1$ and f_2 that satisfy the hypotheses of Theorem 3. Firstly, using the properties of strong typical and conditional typical projectors [39, Properties 15.2.4 and 15.2.7] we have the first four inequalities (hypotheses (7a)) satisfied for all $\epsilon_1, \epsilon_2 \in (0,1)$, and for all sufficiently large n. Next, using [39, Property 15.1.2], there exist functions $\delta_1(\delta), \delta_2(\delta) \searrow 0$ as $\delta \searrow 0$, such that for all sufficiently large n, the first two inequalities of hypotheses (7b) are satisfied for $D_1 \triangleq 2^{n(S(RB)_{\sigma_1}+\delta_1(\delta))}$ and $D_2 \triangleq 2^{n(S(RA)_{\sigma_2}+\delta_2(\delta))}$. Further, using [39, Property 15.2.3], there exist functions $\delta_3(\delta), \delta_4(\delta) \searrow 0$ as $\delta \searrow 0$, such that for all sufficiently large n, the next two inequalities of hypotheses (7b) are satisfied for $d_1 \triangleq 2^{n(S(RB|U)_{\sigma_1}-\delta_3(\delta))}$, $d_2 \triangleq 2^{n(S(RA|V)_{\sigma_2}-\delta_4(\delta))}$. The next four inequalities of hypotheses (7c) follow from the definition of projectors $\Pi_{\rho_A,\delta}$, $\Pi_{\rho_B,\delta}$, $\Pi_{u^n,\delta}^A$ and $\Pi_{v^n,\delta}^B$. And finally, the four inequalities of hypotheses (7d) are satisfied by using [39, Property 15.1.3] and by defining $F_1 \triangleq 2^{n(S(RB)_{\sigma_1}-\delta_1(\delta))}$, $F_2 \triangleq 2^{n(S(RA)_{\sigma_2}-\delta_2(\delta))}$, and $f_1 \triangleq D_1$ and $f_2 \triangleq D_2$.

This implies the existence of a distributed protocol with parameters (n, T_1, T_2, N_1, N_2) with $\Xi_{\rho_{AB}}(M_{AB}, \tilde{M}_{AB}) \leqslant \alpha_A + \alpha_B + \alpha_P$. We now evaluate the upper bound. For this we let $T_i = 2^{nR_i}$, $N_i = 2^{nC_i}$, and $K_i = 2^{n\tilde{R}_i}$, for some non-negative real numbers R_i, C_i, and $\tilde{R}_i$ for $i = 1,2$. Moreover, we assume that $S(U)_{\sigma_3} \geqslant \tilde{R}_1$ and $S(V)_{\sigma_3} \geqslant \tilde{R}_2$. If not, then faithful simulation can be achieved in a trivial way.

Using the property of strongly typical sets, note that for all sufficiently large n we have $|\mathcal{U}| \leqslant 2^{n(S(U)_{\sigma_3}+\delta_5(\delta))}$, $|\mathcal{V}| \leqslant 2^{n(S(V)_{\sigma_3}+\delta_5(\delta))}$, $\lambda_{\mathrm{m}}^A \leqslant 2^{-n(S(U)_{\sigma_3}-\delta_5(\delta))}$ $\lambda_{\mathrm{m}}^B \leqslant 2^{-n(S(V)_{\sigma_3}-\delta_5(\delta))}$. Furthermore, we have the bounds: $|\mathcal{W}| \leqslant 2^{n(S(U,V)_{\sigma_3}+\delta_5(\delta))}$, $W_A \leqslant 2^{n(S(U|V)_{\sigma_3}+\delta_5(\delta))}$, and $W_B \leqslant 2^{n(S(V|U)_{\sigma_3}+\delta_5(\delta))}$, where $\delta_5(\delta) \searrow 0$ as $\delta \searrow 0$. For all sufficiently large n we have $\theta_i \leqslant \epsilon_i$ for $i = 1,2$. Hence for $i = 1,2$, the term $(2\epsilon_i/(1+\epsilon_i)) + f(\epsilon_i, \theta_i)$ can be made arbitrarily small by a suitable choice of ϵ_i and n.

Next we see that

$$\frac{1}{\sqrt{N_1K_1}} \sum_{u\in\mathcal{U}} \sqrt{\lambda_u^A} \leqslant 2^{\left[-\frac{n}{2}(\tilde{R}_1+C_1-S(U)_{\sigma_3}-3\delta_5)\right]}, \text{ and}$$
$$\frac{1}{\sqrt{N_2K_2}} \sum_{v\in\mathcal{V}} \sqrt{\lambda_v^B} \leqslant 2^{\left[-\frac{n}{2}(\tilde{R}_2+C_2-S(V)_{\sigma_3}-3\delta_5)\right]},$$

and hence can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1 + C_1 > S(U)_{\sigma_3} + 3\delta_5 \quad \text{and} \quad \tilde{R}_2 + C_2 > S(V)_{\sigma_3} + 3\delta_5.$$

Moving on, we have

$$\begin{aligned}
&D_1 N_1 \exp\left[-\frac{K_1\epsilon_1^3 d_1 D_1^{-1}}{4\ln 2}\right] \\
&\leqslant 2^{n(S(RB)_{\sigma_1}+C_1+\delta_1)} \exp\left[-\frac{2^{n(\tilde{R}_1-I(RB;U)_{\sigma_1}-\delta_1-\delta_3)}\epsilon_1^3}{4\ln 2}\right],
\end{aligned}$$

[6]Note that $\Pi_{\rho_A,\delta}$ and $\Pi_{\rho_B,\delta}$ also depend on n, however, for ease of notation, we do not make this explicit.

$$D_2 N_2 \exp\left[-\frac{K_2\epsilon_2^3 d_2 D_2^{-1}}{4\ln 2}\right] \leqslant 2^{n(S(RA)_{\sigma_2}+C_2+\delta_2)}\exp\left[-\frac{2^{n(\tilde{R}_2-I(RA;V)_{\sigma_2}-\delta_2-\delta_4)}\epsilon_2^3}{4\ln 2}\right],$$

which can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1 > I(U;RB)_{\sigma_1}+\delta_1+\delta_3, \text{ and } \tilde{R}_2 > I(V;RA)_{\sigma_2}+\delta_2+\delta_4.$$

Next we have

$$\lambda^{AB}(\mathcal{W}^c) = \sum_{(u^n,v^n)\notin T_\delta(U,V)} \lambda^{AB}_{u^n,v^n},$$

which can be made arbitrarily small for sufficiently large n. Finally, we have

$$\left[\frac{\lambda^A_{\text{m}}\lambda^B_{\text{m}}|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2} + \frac{K_1W_A\lambda^A_{\text{m}}}{(1-\theta_1)T_1}\left(1+\frac{\lambda^B_{\text{m}}K_2}{(1-\theta_2)}\right) + \frac{K_2W_B\lambda^B_{\text{m}}}{(1-\theta_2)T_2}\left(1+\frac{\lambda^A_{\text{m}}K_1}{(1-\theta_1)}\right)\right]\frac{f_1f_2}{F_1F_2}$$
$$\leqslant \frac{2^{2n(\delta_1+\delta_2)}}{(1-\theta_1)(1-\theta_2)}\Big[2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_1-R_2-I(U:V)_{\sigma_3}+3\delta_5)\right]} + 2^{\left[n(\tilde{R}_1-R_1-I(U:V)_{\sigma_3}+2\delta_5)\right]} + 2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_1-I(U:V)_{\sigma_3}-S(V)_{\sigma_3}+3\delta_5)\right]} + 2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_2-I(U:V)_{\sigma_3}-S(U)_{\sigma_3}+3\delta_5)\right]} + 2^{\left[n(\tilde{R}_2-R_2-I(U:V)_{\sigma_3}+2\delta_5)\right]}\Big],$$

which again can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1+\tilde{R}_2-R_1-R_2 < I(U;V)_{\sigma_3} - 3\delta_5 - 2(\delta_1+\delta_2).$$

To sum-up, we have showed that the trace distance inequality in (29) holds for all sufficiently small δ, ϵ_1, and ϵ_2, and all sufficiently large n, if the following bounds hold:

$$\begin{aligned} &\tilde{R}_1 > I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 > I(V;RA)_{\sigma_2},\\ &C_1+\tilde{R}_1 > S(U)_{\sigma_3}, \quad C_2+\tilde{R}_2 > S(V)_{\sigma_3},\\ &(\tilde{R}_1-R_1)+(\tilde{R}_2-R_2) < I(U;V)_{\sigma_3},\\ &\tilde{R}_1 \geqslant R_1 \geqslant 0, \quad \tilde{R}_2 \geqslant R_2 \geqslant 0,\\ &C_1 \geqslant 0, \quad C_2 \geqslant 0. \end{aligned} \tag{30}$$

Therefore, there exists a distributed protocol with parameters $(n, 2^{nR_1}, 2^{nR_2}, 2^{nC_1}, 2^{nC_2})$ such that its overall POVM $\tilde{M}^{(n)}_{AB}$ is ϵ-faithful to $M^{\otimes n}_{AB}$ with respect to $\rho^{\otimes n}_{AB}$. Lastly, we complete the proof of the theorem using the following lemma.

Lemma 6: Let $\mathcal{R}_1$ denote the closure of the set of all (R_1,R_2,C_1,C_2) for which there exists $(\tilde{R}_1,\tilde{R}_2)$ such that the sextuple $(R_1,R_2,C_1,C_2,\tilde{R}_1,\tilde{R}_2)$ satisfies the inequalities in (30). Let, $\mathcal{R}_2$ denote the set of all quadruple (R_1,R_2,C_1,C_2) that satisfies the inequalities in (11) given in the statement of the theorem. Then, $\mathcal{R}_1=\mathcal{R}_2$.

Proof: The proof follows by Fourier-Motzkin elimination [44]. □

V. Proofs: Q-C Distributed Rate Distortion Theory

In this section, we provide proofs of the inner and the outer bounds (Theorems 5 and 6) to the achievable rate region of the q-c distributed rate distortion problem.

A. Proof of Theorem 5 (Inner Bound)

In the interest of brevity, we provide the proof for the special case, when the time sharing random variable is trivial, i.e., $\mathcal{Q}$ is empty. An extension to the more general case is straightforward but tedious. For the special case, the proof follows from Theorem 4. Fix POVMs (M_A, M_B) and reconstruction states $S_{u,v}$ as in the statement of the theorem. Let $\mathcal{N}_{AB\mapsto\hat{X}}$ be the mapping corresponding to these POVMs and the reconstruction states. Then, $d(\rho_{AB},\mathcal{N}_{AB\mapsto\hat{X}}) \leqslant D$. According to Theorem 4, for any $\epsilon > 0$, there exists an $(n, 2^{nR_1}, 2^{nR_2}, N_1, N_2)$ distributed protocol for ϵ-faithful simulation of $M_A^{\otimes n}\otimes M_B^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$ such that (R_1,R_2) satisfies the inequalities in (11). Let $\tilde{M}_A^{(\mu_1)}, \tilde{M}_B^{(\mu_2)}, \mu_i \in [1,N_i]$, for $i=1,2$, and $f^{(\mu_1,\mu_2)}$ be the POVMs and the deterministic decoding functions of this protocol with $\mathcal{Z}=\mathcal{U}\times\mathcal{V}$. We use these POVM's and mappings to construct a q-c protocol for distributed quantum source coding.

For each $\mu_i \in [1,N_i]$, for $i=1,2$, consider the q-c protocol with parameters $\Theta_i = 2^{nR_i}, i=1,2$, and POVMs $\tilde{M}_A^{(\mu_1)}, \tilde{M}_B^{(\mu_2)}$. Moreover, we use n-length reconstruction states $S_{i,j} \triangleq \sum_{u^n,v^n}\mathbb{1}\{f^{(\mu_1,\mu_2)}(i,j)=(u^n,v^n)\}S_{u^n,v^n}$, where $S_{u^n,v^n} = \otimes_i S_{u_i,v_i}$. Further, let the corresponding mappings be denoted as $\tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n}$. With this notation, for the average of these random protocols, the following bounds hold:

$$\begin{aligned}
&\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2} d(\rho_{AB}^{\otimes n}, \tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n})\\
&= \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\operatorname{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\\
&= \operatorname{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\\
&\quad + \operatorname{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes(\tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}-\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\\
&\overset{(a)}{\leqslant} \operatorname{Tr}\left\{\Delta\left((\mathrm{id}_R\otimes\mathcal{N}_{AB\mapsto\hat{X}})(\Psi^{\rho_{AB}}_{RAB})\right)\right\}\\
&\quad + \|\Delta^{(n)}(\mathrm{id}\otimes(\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}}-\tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1\\
&\overset{(b)}{\leqslant} D+\|\Delta^{(n)}\|_\infty\|(\mathrm{id}\otimes(\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}}-\tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1\\
&\overset{(c)}{\leqslant} D+\|\Delta^{(n)}\|_\infty\|(\mathrm{id}\otimes(M_A^{\otimes n}\otimes M_B^{\otimes n}-\tilde{M}_{AB}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1\\
&\overset{(d)}{\leqslant} D+\epsilon\|\Delta\|_\infty,
\end{aligned}$$

where $\tilde{\mathcal{N}}_{AB\mapsto\hat{X}}$ is the average of $\tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{AB\mapsto\hat{X}}$, and $\tilde{M}_{AB}$ is the overall POVM of the underlying distributed protocol as given in (6). The inequality (a) holds by the fact that $|\operatorname{Tr}\{A\}| \leqslant ||A||_1$. (b) follows from the fact that for any two operators A and B acting on a Hilbert space $\mathcal{H}$ the following inequalities hold.

$$\|BA\|_1 \leqslant \|B\|_\infty\|A\|_1, \quad \text{and} \quad \|AB\|_1 \leqslant \|B\|_\infty\|A\|_1,$$

(see in [39, Exercise 12.2.1] for a proof). (c) is due to the monotonicity of the trace-distance [39] with respect to the quantum channel given by $\mathrm{id}\otimes\mathcal{L}_{UV\mapsto\hat{X}}^{\otimes n}$, where

$$\mathcal{L}_{UV\mapsto\hat{X}}(\omega)\triangleq\sum_{u,v}\langle u,v|\omega\,|u,v\rangle S_{u,v}.$$

And (d) follows by Theorem 4, and the fact that $\|\Delta^{(n)}\|_\infty\leqslant\|\Delta\|_\infty$. Hence using the collection of codebooks $\{\mathcal{C}^{(\mu_1,\mu_2)}\}_{\mu_1\in[1,N_1],\mu_2\in[1,N_2]}$, constructed in Theorem 4, and averaged over the common randomness, the distortion constraint $\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}d(\rho_{AB}^{\otimes n},\tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}^{(\mu_1,\mu_2)})\leqslant D+\epsilon\|\Delta\|_\infty$ is met. Hence there must exist a realization of the common randomness (μ_1,μ_2), and the corresponding codebook $\mathcal{C}^{(\mu_1,\mu_2)}$ that achieves this distortion. This completes the proof of the theorem, since Δ is a bounded operator.

B. Proof of Theorem 6 (Outer Bound)

Suppose the triplet (R_1,R_2,D) is achievable. Then, from Definition 7, for all $\epsilon>0$, there exists an (n,Θ_1,Θ_2) q-c protocol satisfying the inequalities in the definition. Let $M_A\triangleq\{\Lambda_{l_1}^A\}_{l_1\in[1,\Theta_1]}$, $M_B\triangleq\{\Lambda_{l_2}^B\}_{2\in[1,\Theta_2]}$, and $S_{l_1,l_2}\in\mathcal{D}(\mathcal{H}_{\hat{X}}^{\otimes n})$ be the corresponding POVMs and reconstruction states. Let L_1,L_2 denote the outcomes of the measurements. Then, for Alice's rate, we obtain

$$\begin{aligned}n(R_1+\epsilon)&\geqslant H(L_1)\geqslant H(L_1|L_2)\\&\geqslant I(L_1;R^n|L_2)_\tau=\sum_{j=1}^n I(L_1;R_j|L_2,R^{j-1})_\tau,\end{aligned}$$

where the state τ is defined as $\tau^{L_1L_2R^n\hat{X}^n}\triangleq$

$$\sum_{l_1,l_2}|l_1,l_2\rangle\langle l_1,l_2|\otimes\mathrm{Tr}_{A^nB^n}\left\{(\mathrm{id}\otimes\Lambda_{l_1}^A\otimes\Lambda_{l_2}^B)\Psi_{R^nA^nB^n}^{\rho_{AB}}\right\}\otimes S_{l_1,l_2},$$

and the inequalities follow from L_1 and L_2 being classical. Note that for each j the corresponding mutual information above is defined for a state in the Hilbert space $\mathcal{H}_{L_1}\otimes\mathcal{H}_{L_2}\otimes\mathcal{H}_R^{\otimes j}$. Next, we convert the above summation into a single-letter quantum mutual information term. For that we proceed with defining a new Hilbert space using direct-sum operation.

Let us recall the definition of direct-sum of Hilbert spaces [45]. With this definition, consider the following single-letterization:

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^n I(L_1;R_j|L_2,R^{j-1})_\tau\\&\qquad\overset{(a)}{=}I(L_1;R_J|L_2,R^{J-1},J)_\sigma=I(L_1;R|L_2,Q)_\sigma,\end{aligned}$$

where the state σ is defined as:

$$\begin{aligned}&\sigma^{L_1L_2RQ\hat{X}}\triangleq\\&\sum_{l_1,l_2}\frac{|l_1,l_2\rangle\langle l_1,l_2|}{n}\otimes\Big(\sum_{j=1}^n\big(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\Big\{(\mathrm{id}\\&\quad\otimes\Lambda_{l_1}^A\otimes\Lambda_{l_2}^B)\Psi_{R^nA^nB^n}^{\rho_{AB}}\Big\}\otimes|j\rangle\langle j|\otimes\mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\big)\Big),\end{aligned}\tag{31}$$

and $\mathrm{Tr}_{\hat{X}^n\sim j}$ denotes tracing over $(\hat{X}^{\otimes j-1}\otimes\hat{X}_{j+1}^{\otimes n})$, and $Q\triangleq(R^{J-1},J)$, and J is an averaging random variable which is uniformly distributed over $[1,n]$. We have attached a quantum register for this classical random variable yielding the state σ. The equality (a) follows from the following lemma.

Lemma 7: Consider the classical-quantum state

$$\sigma_{JABC}\triangleq\sum_{j=1}^n P_J(j)\,|j\rangle\langle j|\otimes\rho_{ABC}^j,$$

where $\{|j\rangle\}_{j\in[1,n]}$ is an orthonormal set in some Hilbert space $\mathcal{H}_J$, $\rho_{ABC}^j\in\mathcal{D}(\mathcal{H}_A^j\otimes\mathcal{H}_B^j\otimes\mathcal{H}_C^j)$, where $\{\mathcal{H}_A^j\otimes\mathcal{H}_B^j\otimes\mathcal{H}_C^j\}_{j\in[1,n]}$ is a collection of finite-dimensional Hilbert spaces. Note that $\sigma_{ABC}=\mathrm{Tr}_J(\sigma_{XABC})$ is a state on $\bigoplus_{j=1}^n(\mathcal{H}_A^j\otimes\mathcal{H}_B^j\otimes\mathcal{H}_C^j)$. Then $I(A;B|C,J)_\sigma=\sum_{j=1}^n P_J(j)I(A;B|C)_{\rho^j}$.

Proof: The proof is provided in Appendix B-C. □

We elaborate on the Hilbert space associated with Q as follows. Suppose $\{|\phi_i\rangle\}_{i\in\mathcal{I}}$ is an orthonormal basis for $\mathcal{H}_R$. Then, a basis for $\mathcal{H}_R^{\otimes k}$ is given by

$$|\phi_{\mathbf{i}^k}\rangle\triangleq|\phi_{i_1}\rangle\otimes|\phi_{i_2}\rangle\otimes\cdots\otimes|\phi_{i_k}\rangle,$$

for all $\mathbf{i}^k\in\mathcal{I}^k$. Consider the direct-sum of the Hilbert spaces $\bigoplus_{k=1}^n\mathcal{H}_R^{\otimes k}$ and the Hilbert space $\mathcal{H}_J\otimes\mathcal{H}_R^{\otimes k}$. With this definition, define $\mathcal{H}_Q$, as the Hilbert space which is spanned by $|j\rangle\otimes|\phi_{\mathbf{i}^{(j-1)}}\rangle$, for all $j\in[1,n]$ and $\mathbf{i}^{(j-1)}\in\mathcal{I}^{(j-1)}$. Therefore, $\mathcal{H}_Q$ is *isometrically isomorphic* to the direct-sum $\bigoplus_k\mathcal{H}_R^{\otimes k}$. Note that $\mathcal{H}_Q$ can be viewed as a multi-particle Hilbert space, which is a truncated version of the so-called Fock space [46].

Similarly, for Bob's rate we have

$$R_2+\epsilon\geqslant I(L_2;R|L_1,Q)_\sigma.$$

For the sum-rate, the following inequalities hold

$$\begin{aligned}n(R_1+R_2+2\epsilon)&\geqslant H(L_1,L_2)\geqslant I(L_1,L_2;R^n)_\tau\\&=\sum_{j=1}^n I(L_1,L_2;R_j|R^{j-1})_\tau\\&=nI(L_1,L_2;R|Q)_\sigma,\end{aligned}$$

where the inequalities follow from L_1 and L_2 being classical. In addition, the distortion of this q-c protocol satisfies $d(\rho_{AB}^{\otimes n},\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})\leqslant D+\epsilon$, where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is the quantum channel associated with the protocol. Therefore, as the distortion observable is symbol-wise additive, we obtain

$$\begin{aligned}D+\epsilon&\geqslant\frac{1}{n}\sum_{j=1}^n\mathrm{Tr}\Big\{\Big(\Delta_{R_j\hat{X}_j}\otimes I_{R\hat{X}}^{\otimes[n]\backslash j}\Big)\\&\qquad\times\big(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}\big)\,(\Psi_{R^nA^nB^n}^{\rho_{AB}})\Big\}\\&=\frac{1}{n}\sum_{j=1}^n\mathrm{Tr}\Big\{\Big(\Delta_{R_j\hat{X}_j}\otimes I_{R_1^{j-1}}\otimes I_{R_{j+1}^n\hat{X}^n\sim j}\Big)\\&\qquad\times\big(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}\big)\,(\Psi_{R^nA^nB^n}^{\rho_{AB}})\Big\}\\&=\frac{1}{n}\sum_{j=1}^n\mathrm{Tr}\Big\{\Big(\Delta_{R_j\hat{X}_j}\otimes I_{R_1^{j-1}}\Big)\\&\qquad\times\Big(\mathrm{Tr}_{R_{j+1}^n\hat{X}^n\sim j}\{(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\}\Big)\Big\}\\&\overset{(a)}{=}\mathrm{Tr}\{(\Delta\otimes I_Q)\sigma^{RQ\hat{X}}\},\end{aligned}$$

where (a) holds because of the following argument. From (31), one can show by partially tracing over (L_1, L_2), that

$$\sigma^{RQ\hat{X}} = \mathrm{Tr}_{L_1,L_2}\{\sigma^{L_1L_2RQ\hat{X}}\} = \sum_{j=1}^{n} \frac{1}{n} |j\rangle\langle j| \otimes \mathrm{Tr}_{R_{j+1}^n \hat{X}^n\sim j}\{(\mathrm{id}\otimes \mathcal{N}_{A^nB^n\mapsto \hat{X}^n}) \times (\Psi^{\rho_{AB}}_{R^nA^nB^n})\},$$

and $I_Q \triangleq \sum_{j=1}^{n}\left(I_R^{\otimes(j-1)} \otimes |j\rangle\langle j|\right)$. Then, I_Q is the identity operator acting on $\mathcal{H}_Q$. Therefore, the right-hand side of the equality (a) above can be written as

$$\mathrm{Tr}\{(\Delta \otimes I_Q)\sigma^{RQ\hat{X}}\} = \mathrm{Tr}\left\{\Delta\sigma^{R\hat{X}}\right\}.$$

Let us identify the single-letter quantum test channel as given in the statement of the theorem. First, due to the distributive property of tensor product over direct sum operation, we can rewrite $\sigma^{L_1L_2RQ\hat{X}}$ as

$$\sigma^{L_1L_2RQ\hat{X}} = \Big(\sum_{j=1}^{n}\frac{1}{n}\sum_{l_1,l_2}|l_1,l_2\rangle\langle l_1,l_2| \otimes \Big(\mathrm{Tr}_{R_{j+1}^n A^nB^n}\Big\{(\mathrm{id} \otimes \Lambda^A_{l_1}\otimes\Lambda^B_{l_2})\Psi^{\rho_{AB}}_{R^nA^nB^n}\Big\} \otimes |j\rangle\langle j| \otimes \mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\Big)\Big).$$

Next, we identify a quantum channel $\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}} : \rho_{AB} \mapsto \sigma^{L_1L_2Q\hat{X}}$. For that and for any j define the following intermediate quantum channels:

$$\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}(\omega_{AB}) \triangleq \sum_{l_1,l_2}|l_1,l_2\rangle\langle l_1,l_2| \otimes \Big(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\Big\{(\mathrm{id}_{R^n\sim j}\otimes\Lambda^A_{l_1}\otimes\Lambda^B_{l_2}) \times (\omega_{AB}\otimes E_j)\Big\}\otimes \mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\Big),$$

where $E_j = \Psi^{\rho_{AB}}_{(RAB)^n\sim j}$. One can verify that $\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}$ is indeed a quantum channel. With these definitions, let

$$\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}}(\omega_{AB}) \triangleq \sum_j \frac{1}{n}\left(\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}(\omega_{AB}) \otimes |j\rangle\langle j|\right).$$

Using the property of direct-sum operation, one can verify that $\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}}$ is a valid quantum channel, and moreover,

$$\sigma^{L_1L_2RQ\hat{X}} = (\mathrm{id}\otimes\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}})(\Psi^{\rho_{AB}}_{RAB}).$$

Lastly, we show that the condition $I(R;Q)_\sigma = 0$ is also satisfied. By taking the partial trace of σ over $(L_1, L_2, \hat{X})$ we obtain the following state

$$\begin{aligned}\sigma^{RQ} &= \mathrm{Tr}_{L_1L_2\hat{X}}(\sigma^{L_1L_2RQ\hat{X}})\\ &= \sum_{j=1}^{n}\frac{1}{n}\sum_{l_1,l_2}\Big(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\Big\{(\mathrm{id}\otimes\Lambda^A_{l_1}\otimes\Lambda^B_{l_2}) \times \Psi^{\rho_{AB}}_{R^nA^nB^n}\Big\}\Big)\otimes|j\rangle\langle j|\\ &= \sum_{j=1}^{n}\frac{1}{n}\Big(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\Big\{\Psi^{\rho_{AB}}_{R^nA^nB^n}\Big\}\Big)\otimes|j\rangle\langle j|\\ &= \sum_{j=1}^{n}\frac{1}{n}\Big(\mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\Big)^{\otimes j}\otimes|j\rangle\langle j|\\ &= \mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\otimes\left(\sum_{j=1}^{n}\frac{1}{n}\Big(\mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\Big)^{\otimes(j-1)}\otimes|j\rangle\langle j|\right),\end{aligned}$$

where the last equality is due to the distributive property of tensor product over direct sum operation. Hence, σ^{RQ} is in a tensor product of the form $\sigma^R \otimes \sigma^Q$, and therefore, $I(R;Q)_\sigma = 0$. The proof completes by identifying W_1 and W_2 with L_1 and L_2, respectively.

VI. Simulation of POVMs With Stochastic Processing

Before we provide a proof for our first main result (Theorem 2), we discuss an extension of the Winter's point-to-point measurement compression scheme [1], incorporating additional stochastic processing at the receiver. This extension was first discussed in [2], and also rederived in [14, Corollary 4] and [10]. Since this problem provides us with some of the tools required for the proof of the main result (Theorem 2), developed in Section VII, we rederive its achievability using the approximating POVMs developed in [1]. This will serve as a building block toward proving the main result. In this problem, the receiver (Bob) has access to additional private randomness, and he is allowed to use this additional resource to perform any stochastic mapping of the received classical bits. In fact, the overall effect on the quantum state can be assumed to be a measurement which is a concatenation of the POVM Alice performs and the stochastic map Bob implements. Hence, Alice in this case, does not remain aware of the measurement outcome. It is for this reason that [2] describes this as a non-feedback problem, with the sender not required to know the outcomes of the measurement. With the availability of additional resources, such a formulation is expected to help reduce the overall resources needed.

A. Problem Formulation

Definition 8 (Protocol): For a given finite set $\mathcal{X}$, and a Hilbert space $\mathcal{H}_A$, a measurement simulation protocol with stochastic processing with parameters (n, Θ, N) is characterized by

1) a collection of Alice's sub-POVMs $\tilde{M}^{(\mu)}, \mu \in [1, N]$ each acting on $\mathcal{H}_A^{\otimes n}$ and with outcomes in $[1, \Theta]$, and

2) a collection of Bob's classical stochastic maps $P^{(\mu)}(x^n|l)$ for all $l \in [1, \Theta]$, $x^n \in \mathcal{X}^n$ and $\mu \in [1, N]$.

The overall sub-POVM of this protocol, given by $\tilde{M}$, is characterized by the following operators:

$$\tilde{\Lambda}_{x^n} \triangleq \frac{1}{N}\sum_{\mu,l}P^{(\mu)}(x^n|l)\ \Lambda^{(\mu)}_l, \quad \forall x^n \in \mathcal{X}^n, \tag{32}$$

where $\Lambda^{(\mu)}_l$ are the operators corresponding to the sub-POVMs $\tilde{M}^{(\mu)}$.

In the above definition, Θ characterizes the amount of classical bits communicated from Alice to Bob, and the amount of common randomness is determined by N, with μ being the common randomness bits distributed among the parties. The classical stochastic mappings induced by $P^{(\mu)}$ represents the action of Bob on the received classical bits.

Definition 9 (Achievability): Given a POVM M acting on $\mathcal{H}_A$, and a density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$, a pair (R, C) is said to be achievable, if for all $\epsilon > 0$ and for all sufficiently large n, there exists a measurement simulation protocol with stochastic processing with parameters (n, Θ, N) such that its overall sub-POVM $\tilde{M}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ (see Definition 2), and

$$\frac{1}{n}\log_2 \Theta \leqslant R + \epsilon, \quad \frac{1}{n}\log_2 N \leqslant C + \epsilon.$$

The set of all achievable pairs is called the achievable rate region.

The following theorem characterizes the achievable rate region.

Theorem 8: For any density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$ and any POVM $M \triangleq \{\Lambda_x\}_{x\in\mathcal{X}}$ acting on the Hilbert space $\mathcal{H}_A$, a pair (R, C) is achievable if and only if there exist a POVM $M_A \triangleq \{\Lambda_w^A\}_{w\in\mathcal{W}}$, with $\mathcal{W}$ being a finite set, and a stochastic map $P_{X|W} : \mathcal{W} \to \mathcal{X}$ such that

$$R \geqslant I(R;W)_\sigma \quad \text{and} \quad R + C \geqslant I(RX;W)_\sigma,$$
$$\Lambda_x \triangleq \sum_{w\in\mathcal{W}} P_{X|W}(x|w)\Lambda_w^A, \quad \forall x \in \mathcal{X}.$$

where $\sigma^{RWX} \triangleq \sum_{w,x} \sqrt{\rho}\Lambda_w^A\sqrt{\rho} \otimes P_{X|W}(x|w)\, |w\rangle\langle w| \otimes |x\rangle\langle x|$.

Remark 7: An alternative characterization of the above rate region can also be obtained in terms of Holevo information. For this, we define the following ensemble $\{\lambda_x, \hat{\rho}_x\}$ as

$$\lambda_x = \sum_{w\in\mathcal{W}} \lambda_w^A P_{X|W}(x|w) \quad \text{and} \quad \hat{\rho}_x = \sum_{w\in\mathcal{W}} P_{W|X}(w|x)\hat{\rho}_w^A,$$

for $\{\lambda_w^A, \hat{\rho}_w^A\}$ being the canonical ensemble associated with the POVM M and the state ρ as defined in (5). With this ensemble, we have

$$I(R;W)_\sigma = \chi\left(\{\lambda_w^A, \hat{\rho}_w^A\}\right) \quad \text{and}$$

$$\begin{aligned} I(RX;W)_\sigma &= I(X;W)_\sigma + I(R;XW)_\sigma - I(R;X)_\sigma \\ &= I(X;W)_\sigma + \chi\left(\{\lambda_w^A, \hat{\rho}_w^A\}\right) - \chi\left(\{\lambda_x, \hat{\rho}_x\}\right), \end{aligned}$$

where we have used the Markov Chain $R - W - X$ which is evident from the structure of σ^{RWX}.

As was pointed out in Section I, a proof of achievability and converse for Theorem 8 was provided by Wilde et al. in [2, Section III]. With regards to the proof of achievability, the authors assume [2, Eqns. 53 and 54] to be true, but do not provide a proof for it. Due to the presence of the cut-off operator, which is constructed for the ensemble and not for the individual operators, these equations may not always be true. Since the proof hinges on these two equations and we do not see a straightforward way to prove the two assumptions made (also confirmed in [37]), we provide an alternate proof for achievability below. For the proof of converse, we refer the readers to [2, Section III.3].

B. Proof of Achievability of Theorem 8

Suppose there exist a POVM M_A and a stochastic map $P_{X|W} : \mathcal{W} \to \mathcal{X}$, such that M can be decomposed as

$$\Lambda_x \triangleq \sum_w P_{X|W}(x|w)\Lambda_w^A, \ \forall x \in \mathcal{X}. \tag{33}$$

We begin by defining a canonical ensemble corresponding to M_A as $\{\lambda_w^A, \hat{\rho}_w^A\}_{w\in\mathcal{W}}$. Similarly. for each $w^n \in \mathcal{W}^n$, we also define

$$\tilde{\rho}_{w^n}^A \triangleq \hat{\Pi}\Pi_\rho\Pi_{w^n}\hat{\rho}_{w^n}^A\Pi_{w^n}\Pi_\rho\hat{\Pi},$$

where $\hat{\rho}_{w^n}^A \triangleq \bigotimes_i \hat{\rho}_{w_i}^A$, Π_ρ denotes the δ-typical projector (as in [39, Def. 15.1.3]) corresponding to the density operator ρ, Π_{w^n} denotes the strong conditional typical projector (as in [39, Def. 15.2.4]) corresponding to the canonical ensemble $\{\lambda_w^A, \hat{\rho}_w^A\}_{w\in\mathcal{W}}$, and $\hat{\Pi}$ denotes the projector onto the subspace spanned by the eigenstates of $\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}^A}{(1-\varepsilon)}\Pi_\rho\Pi_{w^n}\hat{\rho}_{w^n}^A\Pi_{w^n}\Pi_\rho$ corresponding to eigenvalues larger than $\varepsilon 2^{-n(S(\rho)+\delta_1)}$, where $\delta_1(\delta)$ is such that $\mathrm{Tr}(\Pi_\rho) \leqslant 2^{n(S(\rho)+\delta_1)}$, and $\varepsilon \triangleq \sum_{w^n\notin\mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}^A$,[7] and $\delta_1 \searrow 0$ as $\delta \searrow 0$.

Using the above definitions, we now construct the approximating POVM.

1) Construction of Random POVMs: In what follows, we construct a collection of random POVMs. Fix R and C as two positive integers. Let $\mu \in [1, 2^{nC}]$ denote the common randomness shared between the sender and receiver. For each $\mu \in [1, 2^{nC}]$, randomly and independently select 2^{nR} sequences $W^{n,(\mu)}(l)$ from the set $\mathcal{W}^n$, according to the pruned distributions, i.e.,

$$\mathbb{P}\left(W^{n,(\mu)}(l) = w^n\right) \triangleq \begin{cases} \dfrac{\lambda_{w^n}^A}{(1-\varepsilon)} & \text{for} \quad w^n \in \mathcal{T}_\delta^{(n)}(W) \\ 0 & \text{otherwise} \end{cases}. \tag{34}$$

Let the collection of operators $\tilde{M}_A^{(n,\mu)}$ be defined as $\{A_{w^n}^{(\mu)} : w^n \in \mathcal{T}_\delta^{(n)}(W)\}$ for each $\mu \in [1, 2^{nC}]$, where $A_{w^n}^{(\mu)}$ is defined as

$$A_{w^n}^{(\mu)} \triangleq \gamma_{w^n}^{(\mu)}\left(\sqrt{\rho}^{-1}\tilde{\rho}_{w^n}^A\sqrt{\rho}^{-1}\right) \quad \text{and}$$
$$\gamma_{w^n}^{(\mu)} \triangleq \frac{1}{2^{nR}}\sum_{l=1}^{2^{nR}} \frac{(1-\varepsilon)}{(1+\eta)}\mathbb{1}_{\{W^{n,(\mu)}(l)=w^n\}}, \tag{35}$$

with $\eta \in (0, 1)$ determining the probability that $\tilde{M}_A^{(n,\mu)}$ does not form a sub-POVM, for all $\mu \in [1, 2^{nC}]$. Since the construction is very similar to the one used in Section IV-B and IV-C, we make a claim similar to the one in Lemma 2 (also see Proposition 8). This claim gives us the first constraint on the classical rate of communication R, which ensures that the operators constructed above for all $\mu \in [1, 2^{nC}]$ are valid sub-POVMs with high probability. Let $\mathbb{1}_{\{\mathrm{sP}\}}$ denote the indicator random variable corresponding to this event. The claim is as follows. For any $\epsilon \in (0, 1)$, $\eta \in (0, 1)$, any $\delta \in (0, 1)$

[7]Note that Π_ρ, Π_{w^n} and $\hat{\Pi}$ depend on n, and δ however, for ease of notation, we do not make this explicit.

sufficiently small, and any n sufficiently large, we have $\mathbb{E}[\mathbb{1}_{\{\text{sP}\}}] \geqslant (1-\epsilon)$ if $R > I(R;W)_\sigma$, where the definition of σ_{RWX} follows from the statement of theorem. From this, let $[\tilde{M}_A^{(n,\mu)}]$ denote the completion of the corresponding sub-POVM $\tilde{M}_A^{(n,\mu)}$ for $\mu \in [1, 2^{nC}]$. Let the operators completing these POVMs, given by $I - \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} A_{w^n}^{(\mu)}$, be denoted by $A_{w_0^n}^{(\mu)}$ for some $w_0^n \notin \mathcal{T}_\delta^{(n)}(W)$, for all $\mu \in [1, 2^{nC}]$, and $A_{w^n}^{(\mu)} = 0$ for $w^n \notin \mathcal{T}_\delta^{(n)}(W) \bigcup \{w_0^n\}$. We use the trivial POVM $\{I\}$ in the case of the complementary event that the operators do not form sub-POVMs for all μ, and associate it with the sequence $\{w_0^n\}$. The POVM is given by $\{\mathbb{1}_{\{\text{sP}\}} A_{w^n}^{(\mu)} + (1-\mathbb{1}_{\{\text{sP}\}})\mathbb{1}_{\{w^n = w_0^n\}} I\}_{w^n \in \mathcal{W}^n}$. Using this construction, we define the intermediate approximating POVM $\tilde{M}_A^{(n)}$ as $\tilde{M}_A^{(n)} = \frac{1}{2^{nC}} \sum_\mu \tilde{M}_A^{(n,\mu)}$ and the operators of $\tilde{M}_A^{(n)}$ as

$$\tilde{\Lambda}_{w^n}^A \triangleq \left(\frac{1}{2^{nC}} \sum_\mu A_{w^n}^{(\mu)} \right) \mathbb{1}_{\{\text{sP}\}} + (1 - \mathbb{1}_{\{\text{sP}\}}) \mathbb{1}_{\{w^n = w_0^n\}} I.$$

Now, we define Bob's stochastic map as $P_{X|W}^n$, yielding the operators of the final approximating POVM as

$$\sum_{w^n \in \mathcal{W}^n} P_{X|W}^n(x^n|w^n) \tilde{\Lambda}_{w^n}^A, \quad x^n \in \mathcal{X}^n.$$

2) Trace Distance: Fix an arbitrary $\epsilon \in (0,1)$. Now, we compare the action of this approximating POVM on the input state $\rho^{\otimes n}$ with that of the given POVM $M^{\otimes n}$, using the characterization provided in Definition 2. Specifically, we show using the expressions for canonical ensemble that, under certain conditions on (R, C), for all sufficiently large n we have $\mathbb{E}[G] \leqslant \epsilon$, where

$$G \triangleq \sum_{x^n \in \mathcal{X}^n} \left\| \sum_{w^n \in \mathcal{W}^n} P_{X|W}^n(x^n|w^n) \sqrt{\rho^{\otimes n}} (\Lambda_{w^n}^A - \tilde{\Lambda}_{w^n}^A) \sqrt{\rho^{\otimes n}} \right\|_1. \tag{36}$$

As a first step, we split and bound G as $G \leqslant S_1 + S_2 + 2(1 - \mathbb{1}_{\{\text{sP}\}})$, where

$$S_1 \triangleq \sum_{x^n} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \right\|_1,$$

$$S_2 \triangleq \sum_{x^n} \left\| P_{X|W}^n(x^n|w_0^n) \frac{1}{2^{nC}} \times \sum_{\mu=1}^{2^{nC}} \left[\sqrt{\rho^{\otimes n}} (I - \sum_{w^n \neq w_0^n} A_{w^n}^{(\mu)}) \sqrt{\rho^{\otimes n}} \right] \right\|_1. \tag{37}$$

Now we bound S_1 by adding and subtracting an appropriate term and using triangle inequality as $S_1 \leqslant S_{11} + S_{12}$, where S_{11} and S_{12} are given by

$$S_{11} \triangleq \left\| \sum_{x^n} \left[\sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| \right] \right\|_1,$$

$$S_{12} \triangleq \left\| \sum_{x^n} \sum_{w^n \neq w_0^n} \left[\frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \right] \otimes |x^n\rangle\langle x^n| \right\|_1.$$

Note that in the above expressions, we have used an additional triangle inequality for block operators (which is in fact an equality) to move the summation over $\mathcal{X}^n$ inside the trace norm. Firstly, we show $\mathbb{E}[S_{11}]$ is small. To simplify the notation, we define $\sigma_{w^n} = \sum_{x^n} P_{X|W}^n(x^n|w^n) |x^n\rangle\langle x^n|$ which gives S_{11} as

$$S_{11} = \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A \otimes \sigma_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\varepsilon)}{(1+\eta)} \sum_{l,\mu} \hat{\rho}_{W^{n,(\mu)}(l)}^A \otimes \sigma_{W^{n,(\mu)}(l)} \right\|_1.$$

We develop the following lemma to bound this term.

Lemma 8: Consider an ensemble given by $\{\tilde{P}_{W^n}(w^n), \mathcal{T}_{w^n}\}$, where $\tilde{P}_{W^n}(w^n)$ is the pruned distribution as defined in (34) and $\mathcal{T}_{w^n}$ is any tensor product state of the form $\mathcal{T}_{w^n} = \bigotimes_{i=1}^n \mathcal{T}_{w_i}$. Then, for any $\epsilon_2 \in (0,1)$, and for all $\eta, \delta \in (0,1)$ sufficiently small, and n sufficiently large, we have

$$\mathbb{E} \left\| \sum_{w^n} \lambda_{w^n}^A \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\varepsilon)}{(1+\eta)} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 \leqslant \epsilon_2, \tag{38}$$

if $R + C > S(\sum_w \lambda_w^A \mathcal{T}_w) - \sum_w \lambda_w^A S(\mathcal{T}_w) = \chi\left(\{\lambda_w^A, \mathcal{T}_w\}\right)$, where $\{W^{n,(\mu)}(l) : l \in [1, 2^{nR}], \mu \in [1, 2^{nC}]\}$ are independent random vectors generated from $\mathcal{W}^n$ according to the pruned distribution given in (34).

Proof: The proof of the lemma is provided in Appendix B-D □

Therefore, using the lemma above with $\mathcal{T}_{w^n} \triangleq \hat{\rho}_{w^n}^A \otimes \sigma_{w^n}$, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[S_{11}] \leqslant \epsilon$ if $R + C > S(\sum_w \lambda_w^A \hat{\rho}_w^A \otimes \sigma_w) - \sum_w \lambda_w^A S(\hat{\rho}_w^A \otimes \sigma_w) = \chi\left(\{\lambda_w^A\}, \{\hat{\rho}_w^A \otimes \sigma_w\}\right) = I(RX;W)_\sigma$, where σ is as defined in the statement of the theorem. Secondly, we bound S_{12} by applying expectation with respect to the codebook generation, and using Gentle Measurement Lemma [39] as follows,

$$\mathbb{E}[S_{12}] \overset{(a)}{\leqslant} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{x^n} \sum_{w^n \neq w_0^n} P_{X|W}^n(x^n|w^n) \mathbb{E}\left[\gamma_{w^n}^{(\mu)} \left\| (\hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A) \right\|_1 \right]$$
$$\overset{(b)}{=} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}^A}{(1+\eta)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1$$

$$= \frac{1}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}^A \left\| \hat{\rho}_{w^n}^A - \hat{\Pi}\Pi_\rho \Pi_{w^n} \hat{\rho}_{w^n}^A \Pi_{w^n} \Pi_\rho \hat{\Pi} \right\|_1$$

$$\overset{(c)}{\leqslant} \frac{(1-\varepsilon)}{(1+\eta)} (2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''}) \triangleq \varepsilon_3, \quad (39)$$

where (a) is obtained by using triangle inequality and the linearity of expectation, (b) is obtained by marginalizing over x^n and using the fact that $\mathbb{E}[\gamma_{w^n}^{(\mu)}] = \frac{\lambda_{w^n}^A}{(1+\eta)}$, and finally (c) uses repeated application of the average gentle measurement lemma, by setting $\varepsilon_3 = \frac{(1-\varepsilon)}{(1+\eta)}(2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''})$ with $\varepsilon_3 \searrow 0$ as $n \to \infty$ for all sufficiently small $\delta > 0$, and, $\varepsilon' \triangleq \varepsilon_p + 2\sqrt{\varepsilon_p}$ and $\varepsilon'' \triangleq 2\varepsilon_p + 2\sqrt{\varepsilon_p}$ for $\varepsilon_p \triangleq 1 - \min\left\{\mathrm{Tr}\,\Pi_\rho \hat{\rho}_{w^n}^A, \mathrm{Tr}\,\Pi_{w^n}\hat{\rho}_{w^n}^A, 1-\varepsilon\right\}$ (see (35) in [2] for details).

Finally, we show that the term corresponding to S_2 can also be made arbitrarily small. This term can be simplified as follows

$$S_2 \leqslant \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{x^n} P_{X|W}^n(x^n|w_0^n) \Bigg\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \sum_{w^n \neq w_0^n} \sqrt{\rho^{\otimes n}} A_{w^n}^{(\mu)} \sqrt{\rho^{\otimes n}} \Bigg\|_1,$$

$$\leqslant \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \Bigg\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A \Bigg\|_1 + \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1$$

$$= S_{21} + S_{22},$$

where

$$S_{21} \triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \Bigg\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \frac{(1-\varepsilon)}{(1+\eta)} \frac{1}{2^{nR}} \sum_{l=1}^{2^{nR}} \hat{\rho}_{W^{n,(\mu)}(l)}^A \Bigg\|_1,$$

$$S_{22} \triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1. \quad (40)$$

Now, for the first term in (40) we use Lemma 8 and claim that for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$, sufficiently small, any n sufficiently large, we have $\mathbb{E}[S_{21}] \leqslant \epsilon$, if

$$R > S\left(\sum_{w \in \mathcal{W}} \lambda_w^A \hat{\rho}_w\right) + \sum_{w \in \mathcal{W}} \lambda_w^A S(\hat{\rho}_w) = I(R;W)_\sigma,$$

where σ is as defined in the statement of the theorem. Note that the requirement we obtain on R was already imposed when claiming the collection of operators $A_{w^n}^{(\mu)}$ forms a sub-POVM. As for the second term in (40) we again use the gentle measurement Lemma and bound its expected value as

$$\mathbb{E}\left[\frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1\right] = \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1+\eta)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1 \leqslant \varepsilon_3,$$

where ε_3 is defined in (39).

In summary, we have performed the following sequence of steps. Firstly, we argued that $\tilde{M}_A^{(n,\mu)}$ forms a valid sub-POVM for all $\mu \in [1, 2^{nC}]$, with high probability, when the rate R satisfies $R > I(R;W)_\sigma$. Secondly, we moved onto bounding the trace norm between the states obtained after the action for these approximating POVMs when compared with those obtained from the action of actual POVM M, characterized as G using Definition 2. As a first step in establishing this bound, we showed that $G \leqslant S_1 + S_2 + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}})$. Firstly, we have shown that $\mathbb{E}[\mathbb{1}_{\{\mathrm{sP}\}}] \geqslant (1-\epsilon)$ if $R > I(R;W)_\sigma$. Then considering S_1, we used the triangle inequality and divided it into two terms: S_{11} and S_{12}. Then, using Lemma 8, we showed that for any given $\epsilon \in (0,1)$, $\mathbb{E}[S_{11}]$ can be made smaller than ϵ, if $R + C > I(RX;W)_\sigma$. As for S_{12}, we showed that it goes to zero in the expected sense using (39). Finally, for the term given by S_2, we bounded this as a sum of two trace norms S_{21} and S_{22} given in (40). We showed that they can be made arbitrarily small in the expected sense if $R > I(R;W)_\sigma$ for all sufficiently large n.

Hence for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large we have $\mathbb{E}[G] \leqslant 6\epsilon$ if

$$R + C > I(RX;W)_\sigma, \quad \text{and} \quad R > I(R;W)_\sigma.$$

Therefore, using random coding arguments, there exists at least one collection of sub-POVMs with the above construction satisfying the statement of Theorem 8.

VII. Proof: Simulation of Distributed POVMs With Stochastic Processing

We provide a proof of Theorem 2 in this section.

A. Construction of an Ensemble of POVMs

Suppose there exist POVMs $M_A \triangleq \{\Lambda_u^A\}_{u \in \mathcal{U}}$ and $M_B \triangleq \{\Lambda_v^B\}_{v \in \mathcal{V}}$ and a stochastic map $P_{Z|UV} : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$, such that M_{AB} can be decomposed as

$$\Lambda_z^{AB} = \sum_{u,v} P_{Z|UV}(z|u,v) \Lambda_u^A \otimes \Lambda_v^B, \ \forall z \in \mathcal{Z}. \quad (41)$$

Note that the proof technique here is very different to the one used in Section IV-C for proving Theorem 4. Recall that in Theorem 4 we initiated the proof by constructing a protocol to faithfully simulate $M_A^{\otimes n} \otimes M_B^{\otimes n}$. However, here we are not interested in faithfully simulating $M_A^{\otimes n} \otimes M_B^{\otimes n}$. Instead, by carefully exploiting the private randomness Charlie possesses, manifested in terms of the stochastic processing applied by him on the classical bits received, i.e., $P_{Z|U,V}$, we aim to strictly reduce the sum rate constraints compared to the ones obtained in (11f) of Theorem 4. This requires a considerably different methodology. More specifically, Lemma 1 was employed in Theorem 4, which guaranteed that any two point-to-point POVMs that can individually approximate their corresponding original POVMs, can also faithfully approximate a measurement formed by the tensor product of the original POVMs performed on any state in the tensor product Hilbert space. Such a lemma cannot be developed in the setting involving a stochastic decoder. This is due to the fact that bits received from Alice and Bob are jointly perturbed by the stochastic decoder which does not allow a straightforward segmentation into two point-to-point problems. The problem becomes analytically tractable using an asymmetric partitioning.

1) Random Coding: We start by generating the canonical ensembles corresponding to M_A and M_B, as given in (5). With this notation, corresponding to each of the probability distributions, we can associate a δ-typical set. Let us denote $\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{T}_\delta^{(n)}(V)$ and $\mathcal{T}_\delta^{(n)}(UV)$ as the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively. Let Π_{ρ_A} and Π_{ρ_B} denote the δ-typical projectors (as in [39, Def. 15.1.3]) for marginal density operators ρ_A and ρ_B, respectively. Also, for any $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$, let $\Pi_{u^n}^A$ and $\Pi_{v^n}^B$ denote the strong conditional typical projectors (as in [39, Def. 15.2.4]) for the canonical ensembles $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$, respectively. For each $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ define

$$\tilde{\rho}_{u^n}^{A'} \triangleq \Pi_{\rho_A}\Pi_{u^n}^A\hat{\rho}_{u^n}^A\Pi_{u^n}^A\Pi_{\rho_A}, \quad \tilde{\rho}_{v^n}^{B'} \triangleq \Pi_{\rho_B}\Pi_{v^n}^B\hat{\rho}_{v^n}^B\Pi_{v^n}^B\Pi_{\rho_B}, \tag{42}$$

where $\hat{\rho}_{u^n}^A \triangleq \bigotimes_i \hat{\rho}_{u_i}^A$ and $\hat{\rho}_{v^n}^B \triangleq \bigotimes_i \hat{\rho}_{v_i}^B$. [8]

With the notation above, define $\sigma^{A'}$ and $\sigma^{B'}$ as

$$\sigma^{A'} \triangleq \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \frac{\lambda_{u^n}^A}{(1-\varepsilon)}\tilde{\rho}_{u^n}^A, \quad \sigma^{B'} \triangleq \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1-\varepsilon')}\tilde{\rho}_{v^n}^B, \tag{43}$$

where $\varepsilon = \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A$ and $\varepsilon' = \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \lambda_{v^n}^B$. Note that $\sigma^{A'}$ and $\sigma^{B'}$ defined above are expectations with respect to the pruned distribution [39]. Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigenstates of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than $\varepsilon 2^{-n(S(\rho_A)+\delta_1)}$ and $\varepsilon' 2^{-n(S(\rho_B)+\delta_1)}$, where $\delta_1 > 0$ is such that $\mathrm{Tr}(\Pi_{\rho_A}) \leqslant 2^{n(S(\rho_A)+\delta_1)}$, and $\mathrm{Tr}(\Pi_{\rho_B}) \leqslant 2^{n(S(\rho_B)+\delta_1)}$, and $\delta_1 \searrow 0$ as $\delta \searrow 0$. Lastly, define

$$\tilde{\rho}_{u^n}^A \triangleq \hat{\Pi}^A\tilde{\rho}_{u^n}^{A'}\hat{\Pi}^A, \quad \text{and} \quad \tilde{\rho}_{v^n}^B \triangleq \hat{\Pi}^B\tilde{\rho}_{v^n}^{B'}\hat{\Pi}^B. \tag{44}$$

In what follows, we construct two random POVMs one for each encoder. Fix a positive integer N and positive real numbers $\tilde{R}_1$ and $\tilde{R}_2$ satisfying $\tilde{R}_1 < S(U)_{\sigma_3}$ and $\tilde{R}_2 < S(V)_{\sigma_3}$, where σ_3 is defined as

$$\sigma_3^{RUV} \triangleq (\mathrm{id}_R \otimes M_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}}),$$

with $\Psi_{RAB}^{\rho_{AB}}$ being any purification of ρ_{AB}. Let $\mu_1 \in [1, N_1]$ denote the common randomness shared between the first encoder and the decoder, and let $\mu_2 \in [1, N_2]$ denote the common randomness shared between the second encoder and the decoder. Let $\tilde{\mu}_1 \in [1, \tilde{N}_1]$ and $\tilde{\mu}_2 \in [1, \tilde{N}_2]$ denote additional pairwise shared randomness used for random coding purposes. This randomness is only used to show the existence of a desired distributed protocol (as defined in Definition 3), and is used only for bounding purposes. We denote $\bar{\mu}_i \triangleq (\mu_i, \tilde{\mu}_i)$, and $\bar{N}_i \triangleq N_i \cdot \tilde{N}_i$ for $i = 1, 2$. For each $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, randomly and independently select $2^{n\tilde{R}_1}$ and $2^{n\tilde{R}_2}$ sequences $(U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k))$ according to the pruned distributions, i.e.,

$$\mathbb{P}\left((U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k)) = (u^n, v^n)\right) = \begin{cases} \dfrac{\lambda_{u^n}^A}{(1-\varepsilon)}\dfrac{\lambda_{v^n}^B}{(1-\varepsilon')} & \text{for} \quad u^n \in \mathcal{T}_\delta^{(n)}(U), v^n \in \mathcal{T}_\delta^{(n)}(V) \\ 0 & \text{otherwise} \end{cases}. \tag{45}$$

Let $\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}$ denote the codebook containing all pairs of codewords $(U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k))$. Construct operators

$$A_{u^n}^{(\bar{\mu}_1)} \triangleq \gamma_{u^n}^{(\bar{\mu}_1)}\left(\sqrt{\rho_A}^{-1}\tilde{\rho}_{u^n}^A\sqrt{\rho_A}^{-1}\right) \quad \text{and}$$
$$B_{v^n}^{(\bar{\mu}_2)} \triangleq \zeta_{v^n}^{(\bar{\mu}_2)}\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^B\sqrt{\rho_B}^{-1}\right), \tag{46}$$

where

$$\gamma_{u^n}^{(\bar{\mu}_1)} \triangleq \frac{1-\varepsilon}{1+\eta}2^{-n\tilde{R}_1}|\{l : U^{n,(\bar{\mu}_1)}(l) = u^n\}| \quad \text{and}$$
$$\zeta_{v^n}^{(\bar{\mu}_2)} \triangleq \frac{1-\varepsilon'}{1+\eta}2^{-n\tilde{R}_2}|\{k : V^{n,(\bar{\mu}_2)}(k) = v^n\}|, \tag{47}$$

where $\eta \in (0,1)$ is a parameter that determines the probability of not obtaining sub-POVMs. Then, for each $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, construct $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ as in the following

$$M_1^{(n,\bar{\mu}_1)} \triangleq \{A_{u^n}^{(\bar{\mu}_1)} : u^n \in \mathcal{T}_\delta^{(n)}(U)\}, \quad \text{and}$$
$$M_2^{(n,\bar{\mu}_2)} \triangleq \{B_{v^n}^{(\bar{\mu}_2)} : v^n \in \mathcal{T}_\delta^{(n)}(V)\}. \tag{48}$$

We show later that $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ form sub-POVMs, with high probability, for all $\bar{\mu} \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, respectively. These collections $\tilde{M}_1^{(n,\bar{\mu}_1)}$ and $\tilde{M}_2^{(n,\bar{\mu}_2)}$ are completed using the operators $I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}$ and $I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}$, and these operators are associated with sequences u_0^n and v_0^n, which are chosen arbitrarily from $\mathcal{U}^n \backslash \mathcal{T}_\delta^{(n)}(U)$ and $\mathcal{V}^n \backslash \mathcal{T}_\delta^{(n)}(V)$, respectively. For $(\tilde{\mu}_1, \tilde{\mu}_2) \in [1, \tilde{N}_1] \times [1, \tilde{N}_2]$, let $\mathbb{1}_{\{\text{sP-i}\}}(\tilde{\mu}_1, \tilde{\mu}_2)$ denote the indicator random variable corresponding to the event that $M_i^{(n,\mu_i,\tilde{\mu}_i)}$ form sub-POVM for all $\mu_i \in [1, N_i]$ for $i = 1, 2$. We use the trivial POVM $\{I\}$ in the case of the complementary event and associate it with u_0^n and v_0^n as the case maybe. In summary, the POVMs are given by $\{\mathbb{1}_{\{\text{sP-1}\}}A_{u^n}^{(\bar{\mu}_1)} + (1 - \mathbb{1}_{\{\text{sP-1}\}})\mathbb{1}_{\{u^n = u_0^n\}}I\}_{u^n \in \mathcal{U}^n}$, and $\{\mathbb{1}_{\{\text{sP-2}\}}B_{v^n}^{(\bar{\mu}_2)} + (1 - \mathbb{1}_{\{\text{sP-2}\}})\mathbb{1}_{\{v^n = v_0^n\}}I\}_{v^n \in \mathcal{V}^n}$.

2) Binning of POVMs: Fix binning rates (R_1, R_2) and choose a $(\bar{\mu}_1, \bar{\mu}_2)$ pair. For each sequence $u^n \in \mathcal{T}_\delta^{(n)}(U)$ assign an index from $[1, 2^{nR_1}]$ randomly and uniformly, such that the assignments for different sequences are done independently. Perform a similar random and independent assignment for all $v^n \in \mathcal{T}_\delta^{(n)}(V)$ with indices chosen from $[1, 2^{nR_2}]$. Repeat this assignment for every $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$. For each $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$, let $\mathcal{B}_1^{(\bar{\mu}_1)}(i)$ and $\mathcal{B}_2^{(\bar{\mu}_2)}(j)$ denote the i^{th} and the j^{th} bins, respectively. More precisely, $\mathcal{B}_1^{(\bar{\mu}_1)}(i)$ is the set of all u^n sequences with assigned index equal to i, and similar is $\mathcal{B}_2^{(\bar{\mu}_2)}(j)$. Moreover

[8] Note that $\tilde{\rho}_{u^n}^A$ and $\tilde{\rho}_{v^n}^B$ are not tensor products operators.

let $\iota_1^{(\bar{\mu}_1)}: \mathcal{T}_\delta^{(n)}(U) \to [1, 2^{nR_1}]$, and $\iota_2^{(\bar{\mu}_2)}: \mathcal{T}_\delta^{(n)}(V) \to [1, 2^{nR_2}]$, denote the corresponding random binning functions. Define the following operators:

$$\Gamma_i^{A,(\bar{\mu}_1)} \triangleq \sum_{u^n \in \mathcal{B}_1^{(\bar{\mu}_1)}(i)} A_{u^n}^{(\bar{\mu}_1)}, \quad \text{and} \quad \Gamma_j^{B,(\bar{\mu}_2)} \triangleq \sum_{v^n \in \mathcal{B}_2^{(\bar{\mu}_2)}(j)} B_{v^n}^{(\bar{\mu}_2)},$$

for all $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$. Using these operators, we form the following collections:

$$M_A^{(n,\bar{\mu}_1)} \triangleq \{\Gamma_i^{A,(\bar{\mu}_1)}\}_{i\in[1,2^{nR_1}]},\ M_B^{(n,\bar{\mu}_2)} \triangleq \{\Gamma_j^{B,(\bar{\mu}_2)}\}_{j\in[1,2^{nR_2}]}.$$

Note that if $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ are sub-POVMs, then so are $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$. This is due to the relations

$$\sum_i \Gamma_i^{A,(\bar{\mu}_1)} = \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}, \quad \sum_j \Gamma_j^{B,(\bar{\mu}_2)} = \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}.$$

To make $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$ complete, we define $\Gamma_0^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)}$ as $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)}$, respectively.[9] In the event that the operators do not form sub-POVM, the sequence u_0^n and v_0^n are mapped to 0. Now, we intend to use the completions $[M_A^{(n,\bar{\mu}_1)}]$ and $[M_B^{(n,\bar{\mu}_2)}]$ as the POVMs for each encoder. Also, note that the effect of the binning is in reducing the communication rates from $(\tilde{R}_1, \tilde{R}_2)$ to (R_1, R_2).

3) Decoder Mapping: We define a mapping $F^{(\bar{\mu}_1,\bar{\mu}_2)}$ acting on the outputs of $[M_A^{(n,\bar{\mu}_1)}]\otimes[M_B^{(n,\bar{\mu}_2)}]$ as follows. On observing $(\bar{\mu}_1, \bar{\mu}_2)$, and the classical indices $(i,j) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ communicated by the encoders, the decoder creates a set as follows:

$$D_{i,j}^{(\bar{\mu}_1,\bar{\mu}_2)} \triangleq \Big\{(u^n, v^n) \in \mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)} : (u^n, v^n) \in \mathcal{T}_\delta^{(n)}(UV) \text{ and } (u^n, v^n) \in \mathcal{B}_1^{(\bar{\mu}_1)}(i) \times \mathcal{B}_2^{(\bar{\mu}_2)}(j)\Big\}.$$

For every $\bar{\mu}_i \in [1 : \bar{N}_i]$, $i \in [1 : 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$ define the function $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u^n, v^n)$ if (u^n, v^n) is the only element of $D_{i,j}^{(\bar{\mu}_1,\bar{\mu}_2)}$; otherwise $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u_0^n, v_0^n)$ Further, $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u_0^n, v_0^n)$ for $i = 0$ or $j = 0$. Finally, the decoder produces $z^n \in \mathcal{Z}^n$ according to the stochastic map $P_{Z|UV}^n(z^n | F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j))$. With this mapping, we form the following collections of operators, for every $(\tilde{\mu}_1, \tilde{\mu}_2)$,

$$\begin{aligned}&\tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) \\ &\triangleq \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1=1}^{N_1} \sum_{\mu_2=1}^{N_2} \sum_{(i,j): F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)=(u^n,v^n)} \Big(\Gamma_i^{A,(\bar{\mu}_1)} \\ &\otimes \Gamma_j^{B,(\bar{\mu}_2)}\Big) + (1 - \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}})(I \otimes I) \mathbb{1}_{\{(u^n,v^n)=(u_0^n,v_0^n)\}},\end{aligned}$$

for all $(u^n, v^n) \in \mathcal{U}^n \times \mathcal{V}^n$. Note that for $\tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) = 0$ for $(u^n, v^n) \notin (\mathcal{T}_\delta^{(n)}(U) \times \mathcal{T}_\delta^{(n)}(V)) \bigcup \{(u_0^n, v_0^n)\}$. We use the stochastic mapping to define the approximating sub-POVM $\tilde{M}_{AB}^{(n)}(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \{\hat{\Lambda}_{z^n}(\tilde{\mu}_1, \tilde{\mu}_2)\}$ as

$$\hat{\Lambda}_{z^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \sum_{u^n,v^n} \tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) P_{Z|U,V}^n(z^n | u^n, v^n),$$

$\forall z^n \in \mathcal{Z}^n$. The performance of the above ensemble is bounded from above as

$$\begin{aligned}&\frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1,\tilde{\mu}_2} \Xi_{\rho_{AB}^{\otimes n}}(M_{AB}^{\otimes n}, \tilde{M}_{AB}^{(n)}(\tilde{\mu}_1, \tilde{\mu}_2)) \\ &\leqslant \frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1,\tilde{\mu}_2} \Big[G(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1, \tilde{\mu}_2)) \mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \\ &\qquad + 2(1 - \mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1, \tilde{\mu}_2))\Big],\end{aligned}$$

where

$$\begin{aligned}&G(\tilde{\mu}_1, \tilde{\mu}_2) \\ &\triangleq \sum_{z^n} \Big\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \Big(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B P_{Z|U,V}^n(z^n|u^n, v^n) \\ &\qquad - \tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) P_{Z|U,V}^n(u^n, v^n)\Big) \sqrt{\rho_{AB}^{\otimes n}} \Big\|_1. \quad (49)\end{aligned}$$

In what follows, under the conditions on the rates given in the theorem, we show the existence of a pair $(\tilde{\mu}_1, \tilde{\mu}_2)$, and codebooks $\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}$ and binning functions $\iota_i^{(\bar{\mu}_i)}$, for $\mu_i \in [1, N_i]$, $i = 1, 2$, such that the ϵ-faithfulness is satisfied for an arbitrary $\epsilon > 0$ for all sufficiently large n.

B. Performance Analysis

Step 0 (Operators Form Sub-POVM): Fix an arbitrary $\epsilon > 0$. To start with, for all $(\tilde{\mu}_1, \tilde{\mu}_2) \in [1, \tilde{N}_1] \times [1, \tilde{N}_2]$, one can show using a result similar to Lemma 2 the following proposition.

Proposition 3 (sub-POVM): For any $\epsilon \in (0,1)$, any $\eta \in (0,1)$, any $\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have

$$\frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1,\tilde{\mu}_2} 2\Big(1 - \mathbb{E}\Big[\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1, \tilde{\mu}_2)\Big]\Big) < 2\epsilon,$$

if $\tilde{R}_1 > I(U; RB)_{\sigma_1}$ and $\tilde{R}_2 > I(V; RA)_{\sigma_2}$, where σ_1, σ_2 are defined as in the statement of the theorem.

Proof: We skip the proof for brevity. □

Next we focus on G.

Step 1 (Isolating the Effect of Error Induced by Not Covering: Consider the second term within $G(\tilde{\mu}_1, \tilde{\mu}_2)$, which, under the event $\mathbb{1}_{\{\text{sP-1}\}} = 1$ and $\mathbb{1}_{\{\text{sP-2}\}} = 1$, can be written as

$$\begin{aligned}&\sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2)) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(u^n, v^n) \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{i,j} \sqrt{\rho_{AB}^{\otimes n}} \Big(\Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)}\Big) \sqrt{\rho_{AB}^{\otimes n}} \\ &\quad \times P_{Z|U,V}^n(z^n | F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)) \underbrace{\sum_{u^n,v^n} \mathbb{1}_{\{F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)=(u^n,v^n)\}}}_{=1} \\ &= T(\tilde{\mu}_1, \tilde{\mu}_2) + \widetilde{T}(\tilde{\mu}_1, \tilde{\mu}_2),\end{aligned}$$

[9]Note that $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)} = I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)} = I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}$.

where

$$T(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i,j>0}\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_i^{A,(\bar{\mu}_1)}\otimes\Gamma_j^{B,(\bar{\mu}_2)}\right) \times \sqrt{\rho_{AB}^{\otimes n}}P^n_{Z|U,V}(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)),$$

$$\widetilde{T}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i=0 \text{ or } j=0}\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_i^{A,(\bar{\mu}_1)}\otimes\Gamma_j^{B,(\bar{\mu}_2)}\right) \times \sqrt{\rho_{AB}^{\otimes n}}P^n_{Z|U,V}(z^n|u_0^n,v_0^n).$$

Hence, we have

$$G(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} \leqslant [S(\tilde{\mu}_1,\tilde{\mu}_2)+\widetilde{S}(\tilde{\mu}_1,\tilde{\mu}_2)]\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}, \quad (50)$$

where $S(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq$

$$\sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda^A_{u^n}\otimes\Lambda^B_{v^n}P^n_{Z|U,V}(z^n|u^n,v^n)\right) \times\sqrt{\rho_{AB}^{\otimes n}} - T(\tilde{\mu}_1,\tilde{\mu}_2)\right\|_1, \quad (51)$$

and $\widetilde{S}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \sum_{z^n}\|\widetilde{T}(\tilde{\mu}_1,\tilde{\mu}_2)\|_1$. Note that $\widetilde{S}$ captures the error induced by not covering the state $\rho_{AB}^{\otimes n}$.

Remark 8: The terms corresponding to the operators that complete the sub-POVMs $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$, i.e., $I-\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}A_{u^n}^{(\bar{\mu}_1)}$ and $I-\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}B_{v^n}^{(\bar{\mu}_2)}$ are taken care of in $\widetilde{T}$. The expression T excludes the completing operators. Therefore, in the analysis of the term S, we use $A_{u^n}^{(\bar{\mu}_1)}$ and $B_{v^n}^{(\bar{\mu}_2)}$ to denote the operators corresponding to $u^n\in\mathcal{T}_\delta^{(n)}(U)$ and $v^n\in\mathcal{T}_\delta^{(n)}(V)$, respectively.

Step 2 (Isolating the Effect of Error Induced by Binning): Noting that $e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)=F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)$, for each $(u^n,v^n)\in\mathcal{B}_1^{(\bar{\mu}_1)}(i)\times\mathcal{B}_2^{(\bar{\mu}_2)}(j)$ and $(u^n,v^n)\in\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}$. For any $(u^n,v^n)\notin\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}$ let $e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)=(u_0^n,v_0^n)$. This simplifies T as

$$\begin{aligned}T(\tilde{\mu}_1,\tilde{\mu}_2) &= \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{\substack{i>0,\\ j>0}}\sqrt{\rho_{AB}^{\otimes n}}\Bigg(\sum_{u^n\in B_1^{(\bar{\mu}_1)}(i)}A_{u^n}^{(\bar{\mu}_1)} \otimes\sum_{v^n\in B_2^{(\bar{\mu}_2)}(j)}B_{v^n}^{(\bar{\mu}_2)}\Bigg)\sqrt{\rho_{AB}^{\otimes n}}P^n_{Z|U,V}(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j))\\ &=\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}} \times\sum_{\substack{i>0,\\ j>0}}\mathbb{1}_{\left\{u^n\in B_1^{(\bar{\mu}_1)}(i),v^n\in B_2^{(\bar{\mu}_2)}(j)\right\}}P^n_{Z|U,V}(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n))\\ &=\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}} \times P^n_{Z|U,V}(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)),\end{aligned}$$

where we have used the fact that $\sum_{u^n\in B_1^{(\bar{\mu}_1)}(i)}A_{u^n}^{(\bar{\mu}_1)} = \sum_{u^n}A_{u^n}^{(\bar{\mu}_1)}\mathbb{1}_{\left\{u^n\in B_1^{(\bar{\mu}_1)}(i)\right\}}$ and $\sum_{i>0}\mathbb{1}_{\left\{u^n\in B_1^{(\bar{\mu}_1)}(i)\right\}}=1$ for all $u^n\in\mathcal{T}_\delta^{(n)}(U)$, and a similar argument holds for the sub-POVM $\{B_{v^n}^{(\bar{\mu}_2)}\}$. Note that the (u^n,v^n) that appear in the above summation is confined to $(\mathcal{T}_\delta^{(n)}(U)\times\mathcal{T}_\delta^{(n)}(V))$, however for ease of notation, we do not make this explicit. We substitute the above expression into S as in (51) to obtain

$$\begin{aligned}&S(\tilde{\mu}_1,\tilde{\mu}_2)\\ &=\sum_{z^n}\Bigg\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda^A_{u^n}\otimes\Lambda^B_{v^n}\right)\sqrt{\rho_{AB}^{\otimes n}}P^n_{Z|U,V}(z^n|u^n,v^n) -\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}} \times P^n_{Z|U,V}(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n))\Bigg\|_1.\end{aligned}$$

Recall that $\bar{\mu}_i=(\mu_i,\tilde{\mu}_i)$ for $i=1,2$. We add and subtract an appropriate term within S and apply triangle inequality to isolate the effect of binning as $S\leqslant S_1+S_2$, where

$$\begin{aligned}&S_1(\tilde{\mu}_1,\tilde{\mu}_2)\\ &\triangleq\sum_{z^n}\Bigg\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\Bigg(\Lambda^A_{u^n}\otimes\Lambda^B_{v^n}-\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\Bigg) \times\sqrt{\rho_{AB}^{\otimes n}}P^n_{Z|U,V}(z^n|u^n,v^n)\Bigg\|_1,\\ &S_2(\tilde{\mu}_1,\tilde{\mu}_2)\\ &\triangleq\sum_{z^n}\Bigg\|\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}} \times\left(P^n_{Z|U,V}(z^n|u^n,v^n)-P^n_{Z|U,V}\left(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\right)\right)\Bigg\|_1.\end{aligned} \quad (52)$$

This gives

$$G\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\leqslant[S_1+S_2+\widetilde{S}]\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}.$$

Note that the term S_1 characterizes the error introduced by approximation of the original POVM with the collection of approximating sub-POVMs $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$, and the term S_2 characterizes the error caused by binning of these approximating sub-POVMs. Next, we analyze S_2 and prove the following proposition.

Proposition 4 (Mutual Packing): For any $\epsilon\in(0,1)$, any $\eta,\delta\in(0,1)$ sufficiently small, and any n sufficiently large n, we have

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}\mathbb{E}\Big[S_2(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\Big]<5\epsilon,$$

if $\tilde{R}_1>I(U;RB)_{\sigma_1}$, $\tilde{R}_2>I(V;RA)_{\sigma_2}$, $\tilde{R}_1+\frac{1}{n}\log(\bar{N}_1)>S(U)_{\sigma_3}$, $\tilde{R}_2+\frac{1}{n}\log(\bar{N}_2)>S(V)_{\sigma_3}$, $\tilde{R}_1+\tilde{R}_2-R_1-R_2<I(U;V)_{\sigma_3}$, where σ_i for $i=1,2,3$, is the auxiliary state defined in the theorem.

Proof: The proof is provided in Appendix C-C □

Hence there must exist a pair $(\tilde{\mu}_1,\tilde{\mu}_2)$ such that

$$2\left(1-\mathbb{E}\left[\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2))\right]\right) +\mathbb{E}\left[S_2(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\right]<7\epsilon,$$

for the rates satisfying the constraints in Propositions 3 and 4. For the rest of the proof, we fix $(\tilde{\mu}_1,\tilde{\mu}_2)$ to be this pair.

The dependence of functions defined in the sequel on this pair is not made explicit for ease of notation.

Remark 9: Since the shared randomness given by $(\tilde{\mu}_1, \tilde{\mu}_2)$ is only used for random coding purposes, two of the constraints in Proposition 4, given by $\tilde{R}_1 + \frac{1}{n}\log(\bar{N}_1) > S(U)_{\sigma_3}$, $\tilde{R}_2 + \frac{1}{n}\log(\bar{N}_2) > S(V)_{\sigma_3}$, are superfluous.

For the term corresponding to $\widetilde{S}$, we prove the following result.

Proposition 5: For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have

$$\mathbb{E}\left[\widetilde{S}\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right] < 8\epsilon,$$

if $\tilde{R}_1 > I(U;RB)_{\sigma_1}$ and $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, where σ_1 and σ_2 are auxiliary states defined in the theorem.

Proof: The proof is provided in Appendix C-D. □

Step 3 (Isolating the Effect of Alice's Approximating Measurement): In this step, we separately analyze the effect of approximating measurements at the two distributed parties in the term S_1. For that, we split S_1 as $S_1 \leqslant Q_1 + Q_2$, where

$$Q_1 \triangleq \sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \frac{1}{N_1}\sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1,$$

$$Q_2 \triangleq \sum_{z^n}\left\|\frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\mu_1)} \otimes \Lambda_{v^n}^B - \frac{1}{N_2}\sum_{\mu_2=1}^{N_2} A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1.$$

With this partition, the terms within the trace norm of Q_1 differ only in the action of Alice's measurement. And similarly, the terms within the norm of Q_2 differ only in the action of Bob's measurement. Showing that these two terms are small forms a major portion of the achievability proof.

Analysis of Q_1: To show Q_1 is small, we compute rate constraints which ensure that an upper bound to Q_1 can be made to vanish in an expected sense. Furthermore, this upper bound becomes convenient in obtaining a single-letter characterization for the rate needed to make the term corresponding to Q_2 vanish. For this, we define J as

$$J \triangleq \sum_{z^n,v^n}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \frac{1}{N_1}\sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \Lambda_{v^n}^B\right) \times \sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1. \quad (53)$$

By defining J and using triangle inequality for block operators (which holds with equality), we add the sub-system V to RZ, resulting in the joint system RZV, corresponding to the state σ_3 as defined in the theorem. Then we approximate the joint system RZV using an approximating sub-POVM $M_A^{(n)}$ producing outputs on the alphabet $\mathcal{U}^n$. To make J small for all sufficiently large n, we expect the sum of the rate of the approximating sub-POVM and common randomness, i.e., $\tilde{R}_1 + C_1$, to be larger than $I(U;RZV)_{\sigma_3}$. We seek to prove this in the following.

Proposition 6: For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_1] \leqslant \mathbb{E}[J] < 2\epsilon$, if $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$, where the auxiliary state σ_3 is defined in the theorem.

Proof: The proof is provided in Appendix C-E. □

Now we move on to bounding Q_2.

Step 4 (Analyzing the Effect of Bob's Approximating Measurement): Step 3 ensured that the sub-system RZV is close to a tensor product state in trace-norm. In this step, we approximate the state corresponding to the sub-system RZ using the approximating POVM $M_B^{(n)}$, producing outputs on the alphabet $\mathcal{V}^n$. We proceed with the following proposition.

Proposition 7 (Non-product Covering Lemma): For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have

$$\mathbb{E}\left[Q_2\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right] < 4\epsilon,$$

if $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$, and $\tilde{R}_2 + C_2 > I(V;RZ)_{\sigma_3}$, where the auxiliary state σ_3 is defined in the theorem.

Proof: The proof is provided in Appendix C-F. □

C. Rate Constraints

To sum-up, we showed that the trace distance satisfies:

$$\Xi_{\rho_{AB}^{\otimes n}}(M_{AB}^{\otimes n}, \tilde{M}_{AB}^{(n)}(\tilde{\mu}_1,\tilde{\mu}_2)) \leqslant 21\epsilon,$$

if the following bounds hold:

$$\begin{gathered}\tilde{R}_1 > I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 > I(V;RA)_{\sigma_2},\\ \tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}, \quad \tilde{R}_2 + C_2 > I(V;RZ)_{\sigma_3},\\ (\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) < I(U;V)_{\sigma_3},\\ \tilde{R}_1 \geqslant R_1 \geqslant 0, \quad \tilde{R}_2 \geqslant R_2 \geqslant 0, \quad C_1 \geqslant 0, \quad C_2 \geqslant 0. \quad (54)\end{gathered}$$

Let us denote the above achievable rate-region by $\mathcal{R}_1$. By doing an exact symmetric analysis, but by replacing the first encoder by a product distribution instead of the second encoder in S_1 (as defined in (52)), all the constraints remain the same, except that the constraints on $\tilde{R}_1 + C_1$ and $\tilde{R}_2 + C_2$ change as follows

$$\tilde{R}_1 + C_1 \geqslant I(U;RZ)_{\sigma_3}, \quad \tilde{R}_2 + C_2 \geqslant I(V;RZU)_{\sigma_3}. \quad (55)$$

Let us denote the above region by $\mathcal{R}_2$. By time sharing between the any two points of $\mathcal{R}_1$ and $\mathcal{R}_2$ one can achieve any point in the convex closure of $(\mathcal{R}_1 \bigcup \mathcal{R}_2)$. The following lemma gives a symmetric characterization of the closure of convex hull of the union of the above achievable rate-regions.

Lemma 9: For the above defined rate regions $\mathcal{R}_1$ and $\mathcal{R}_2$, we have $\mathcal{R}_3 = \text{Convex Closure}(\mathcal{R}_1 \bigcup \mathcal{R}_2)$, where $\mathcal{R}_3$ is given by the set of all the sextuples $(\tilde{R}_1, \tilde{R}_2, R_1, R_2, C_1, C_2)$ satisfying the following constraints:

$$\begin{aligned}&\tilde{R}_1 \geqslant I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 \geqslant I(V;RA)_{\sigma_2},\\ &\tilde{R}_1 + C_1 \geqslant I(U;RZ)_{\sigma_3}, \quad \tilde{R}_2 + C_2 \geqslant I(V;RZ)_{\sigma_3},\\ &\tilde{R}_1 + \tilde{R}_2 + C_1 + C_2 \geqslant I(U;RZ)_{\sigma_3} + I(V;RZ)_{\sigma_3} + I(U;V|RZ)_{\sigma_3},\end{aligned}$$

$$\begin{aligned}\tilde{R}_1+\tilde{R}_2 - (R_1+R_2) &\leqslant I(U;V)_{\sigma_3}\\ 0 \leqslant R_1 \leqslant \tilde{R}_1 \quad 0 \leqslant R_2 &\leqslant \tilde{R}_2 \quad C_1 \geqslant 0, C_2 \geqslant 0. \end{aligned} \tag{56}$$

Proof: The proof follows from elementary convex analysis. □

Lemma 10: Let $\bar{\mathcal{R}}_3$ denote the set of all quadruples (R_1, R_2, C_1, C_2) for which there exists $(\tilde{R}_1, \tilde{R}_2)$ such that the sextuple $(R_1, R_2, C_1, C_2, \tilde{R}_1, \tilde{R}_2)$ satisfies the inequalities in (56). Let $\mathcal{R}_F$ denote the set of all quadruples (R_1, R_2, C_1, C_2) that satisfy the inequalities in (4) given in the statement of the theorem. Then, $\bar{\mathcal{R}}_3 = \mathcal{R}_F$.

Proof: This follows by Fourier-Motzkin elimination [44]. □

VIII. Conclusion

We have developed a distributed measurement compression protocol where we introduced the technique of mutual covering and random binning of distributed measurements. Using these techniques, a set of communication rate-pairs and common randomness rate is characterized for faithful simulation of distributed measurements. We further developed an approach for a distributed quantum-to-classical rate-distortion theory, and provided single-letter inner and outer bounds. As a part of future work, we intend to improve the outer bound by providing a dimensionality bound on the auxiliary Hilbert space involved in the expression. Further, we also desire to improve the achievable rate region by using structured POVMs based on algebraic codes.

Appendix A
Proof of Theorem 7

Note that $\theta = 1 - \sum_{x\in\mathcal{X}} \lambda_x$. Define $\tilde{\rho}'_x \triangleq \Pi_\rho \Pi_x \hat{\rho}_x \Pi_x \Pi_\rho$, and $\sigma' \triangleq \frac{1}{1-\theta}\sum_{x\in\mathcal{X}} \lambda_x \tilde{\rho}'_x$. Further let $\hat{\Pi}$ be the projector onto the subspace spanned by the eigenspace of σ' corresponding to the eigenvalues greater than ϵ/D. Let $\tilde{\rho}_x \triangleq \hat{\Pi}\tilde{\rho}'_x\hat{\Pi}$, and $\sigma \triangleq \hat{\Pi}\sigma'\hat{\Pi}$.

Construction of Random POVMS: Define a collection of random codes $\mathcal{C} \triangleq \{\mathcal{C}^{(\mu)}\}$ for $\mu \in [1, N]$, where $\mathcal{C}^{(\mu)} \triangleq \{X(l,\mu)\}_{l\in[1,K]}$, and $X(l,\mu)$ are chosen randomly, independently according to the distribution $\{\lambda_x/(1-\theta)\}_{x\in\mathcal{X}}$. Using this, define

$$\begin{aligned}\gamma_x^{(\mu)} &\triangleq \frac{(1-\theta)}{(1+\epsilon)}\frac{1}{K}|\{l : X(l,\mu) = x\}|\\ &= \frac{(1-\theta)}{(1+\epsilon)}\frac{1}{K}\sum_{l=1}^{K}\mathbb{1}_{\{X(l,\mu)=x\}},\end{aligned}$$

and $A_x^{(\mu)} \triangleq \gamma_x^{(\mu)}\sqrt{\rho}^{-1}\tilde{\rho}_x\sqrt{\rho}^{-1}$, where $\sqrt{\rho}^{-1}$ refers to the generalized inverse as defined in [38, Section 5.6]. Now for each $\mu \in [1, N]$, construct a collection of non-negative operators $\tilde{M}^{(\mu)} \triangleq \{A_x^{(\mu)}\}_{x\in\mathcal{X}}$.

Proposition 8: $\tilde{M}^{(\mu)}$ forms a sub-POVM for all $\mu \in [1, N]$ with probability exceeding $1 - 2ND\exp\left[-\frac{K\epsilon^2 d\epsilon D^{-1}}{4\ln 2}\right]$.

Proof: We use the operator Chernoff bound [39]. Note that

$$\tilde{\rho}_x \leqslant d^{-1}\hat{\Pi}\Pi_\rho\Pi_x\Pi_\rho\hat{\Pi} \leqslant d^{-1}\hat{\Pi},$$

where we used the hypothesis (15d) assumed in the theorem statement. Moreover,

$$\mathbb{E}[\tilde{\rho}_{X(l,\mu)}] = \hat{\Pi}\sigma'\hat{\Pi} \geqslant \frac{\epsilon}{D}\hat{\Pi}.$$

Applying the operator Chenoff bound on $\{d\tilde{\rho}_{X(l,\mu)}\}_{l\in[1,K]}$, we obtain

$$\begin{aligned}P\left\{(1-\epsilon)\sigma \leqslant \frac{1}{K}\sum_{l=1}^{K}\tilde{\rho}_{X(l,\mu)} \leqslant (1+\epsilon)\sigma\right\}\\ \geqslant 1 - 2D\exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2}\right],\end{aligned}$$

for all $\mu \in [1, N]$, where we used the fact that $\mathrm{Tr}\left(\hat{\Pi}\right) \leqslant \mathrm{Tr}(\Pi_\rho) \leqslant D$ (using the hypothesis (15c) of the theorem statement). Now we have

$$\sigma = \hat{\Pi}\sigma'\hat{\Pi} \leqslant \sigma' \leqslant \frac{1}{1-\theta}\Pi_\rho\rho\Pi_\rho \leqslant \frac{1}{1-\theta}\rho,$$

using the hypothesis (15e) and (15f) of the theorem statement. This results in $(1-\theta)\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} \leqslant I$. This implies that with probability exceeding $1 - 2D\exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2}\right]$, we have

$$\begin{aligned}\sum_{x\in\mathcal{X}} A_x^{(\mu)} &= \sum_{x\in\mathcal{X}}\gamma_x^{(\mu)}\sqrt{\rho}^{-1}\tilde{\rho}_x\sqrt{\rho}^{-1}\\ &= \frac{1}{K}\frac{(1-\theta)}{(1+\epsilon)}\sqrt{\rho}^{-1}\left(\sum_{l=1}^{K}\tilde{\rho}_{X(l,\mu)}\right)\sqrt{\rho}^{-1} \leqslant I.\end{aligned}$$

Hence using the union bound, we see that with probability exceeding $1 - 2ND\exp\left[-\frac{K\epsilon^2 d\epsilon D^{-1}}{4\ln 2}\right]$, we have $\{A_x^{(\mu)}\}_{x\in\mathcal{X}}$ forming a sub-POVM for all $\mu \in [1, N]$. □

Let $\tilde{M} \triangleq \{\tilde{\Lambda}_x\}_{x\in\mathcal{X}}$, $\tilde{M}^{(\mu)} \triangleq \{A_x^{(\mu)}\}_{x\in\mathcal{X}}$, where $\tilde{\Lambda}_x \triangleq \frac{1}{N}\sum_{\mu=1}^{N} A_x^{(\mu)}$. Let $\mathbb{1}_{\{\mathrm{sP}\}}$ denote the indicator random variable corresponding to the event that $\tilde{M}^{(\mu)}$ forms a sub-POVM for all $\mu \in [1, N]$. The completion of the sub-POVM is given by $I - \sum_{x\in\mathcal{X}}\tilde{\Lambda}_x$. We use the trivial POVM $\{I\}$ in the case of the complementary event. Using this construction, we have

$$\begin{aligned}&\Xi_\rho(M, \tilde{M})\\ &\leqslant \mathbb{1}_{\{\mathrm{sP}\}}\left[\sum_{x\in\mathcal{X}}\|\sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho}\|_1 + \mathrm{Tr}\left((I - \sum_{x\in\mathcal{X}}\tilde{\Lambda}_x)\rho\right)\right]\\ &\quad + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}}) + \theta\\ &\leqslant \sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1 + \left\|\rho - \frac{1}{N}\sum_{\mu,x}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1\\ &\quad + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}}) + \theta\\ &\overset{(a)}{\leqslant} \sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1\\ &\quad + \sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1 + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta\\ &\overset{(b)}{\leqslant} 2\sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\hat{\rho}_x\right\|_1\end{aligned}$$

$$+ 2\sum_{x\in\mathcal{X}}\frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\|\hat{\rho}_x-\tilde{\rho}_x\|_1 + 2(1-\mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta$$
$$\overset{(c)}{=} 2[S_1+S_2] + 2(1-\mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta,$$

where (a) follows by triangle inequality, (b) follows by adding and subtracting $\frac{1}{N}\sum_\mu \gamma_x^{(\mu)}\hat{\rho}_x$, and (c) follows by defining

$$S_1 \triangleq \sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\hat{\rho}_x\right\|_1,$$
$$S_2 \triangleq \sum_{x\in\mathcal{X}}\frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\|\hat{\rho}_x-\tilde{\rho}_x\|_1.$$

We work on the first term S_1 as follows. Note that

$$S_1 \leqslant S_1' + \left(\frac{\epsilon}{1+\epsilon}\right)\sum_{x\in\mathcal{X}}\lambda_x \leqslant S_1' + \left(\frac{\epsilon}{1+\epsilon}\right),$$

where

$$S_1' \triangleq \frac{1}{1+\epsilon}\sum_{x\in\mathcal{X}}\left|\lambda_x - \frac{(1-\theta)}{NK}\sum_{\mu,l}\mathbb{1}_{\{X(l,\mu)=x\}}\right|.$$

Note that

$$\mathbb{E}[S_1'] = \frac{1}{(1+\epsilon)}\sum_{x\in\mathcal{X}}\mathbb{E}\left[|\hat{P}(x)-\mathbb{E}[\hat{P}(x)]|\right]$$
$$\leqslant \frac{1}{(1+\epsilon)}\sum_{x\in\mathcal{X}}\sqrt{\mathrm{Var}(\hat{P}(x))} \leqslant \frac{1}{(1+\epsilon)}\sum_{x\in\mathcal{X}}\sqrt{\frac{\lambda_x}{NK}},$$

where we have defined $\hat{P}(x) \triangleq \frac{(1-\theta)}{NK}\sum_{l,\mu}\mathbb{1}_{\{X(l,\mu)=x\}}$. Hence

$$\mathbb{E}[S_1] \leqslant \frac{1}{(1+\epsilon)\sqrt{NK}}\sum_{x\in\mathcal{X}}\sqrt{\lambda_x} + \left(\frac{\epsilon}{1+\epsilon}\right).$$

Moving on to S_2, consider the following.

$$2\mathbb{E}[S_2] \leqslant \frac{2}{(1+\epsilon)}\left[\sum_{x\in\mathcal{X}}\lambda_x\|\hat{\rho}_x-\tilde{\rho}_x'\|_1 + \sum_{x\in\mathcal{X}}\lambda_x\|\tilde{\rho}_x'-\tilde{\rho}_x\|_1\right]$$
$$\leqslant \frac{1}{(1+\epsilon)}\left[4\sqrt{\epsilon}+4\sqrt{\epsilon+2\sqrt{\epsilon}}+4\sqrt{2}(1-\theta)\sqrt{\epsilon+\sqrt{\epsilon}}\right]$$
$$= f(\epsilon,\theta),$$

where we have used the ensemble gentle measurement lemma [39]. Combining all the arguments, we see that

$$\mathbb{E}(\Xi_\rho(M,\tilde{M})) \leqslant \frac{2}{(1+\epsilon)\sqrt{NK}}\sum_{x\in\mathcal{X}}\sqrt{\lambda_x} + \frac{2\epsilon}{\epsilon+1} + f(\epsilon,\theta) + 4DN\exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2}\right] + 2\theta.$$

There must exists a collection of sub-POVMs whose average performance is at least as good.

Appendix B
Proof of Lemmas

A. Proof of Lemma 3

Consider the left hand side of (26). We define an operator Λ_{y_0} which completes the sub-POVM $\{\Lambda_y\}_{y\in\mathcal{Y}}$ as $\Lambda_{y_0} \triangleq I - \sum_{y\in\mathcal{Y}}\Lambda_y$. Further, let the set $\mathcal{Y}^+ \triangleq \mathcal{Y}\bigcup\{y_0\}$. Since trace norm is invariant to transposition with respect to ρ_{AB}, we can write for any $y\in\mathcal{Y}^+$,

$$\left\|\sqrt{\rho_{AB}}(\Gamma^A\otimes\Lambda_y^B)\sqrt{\rho_{AB}}\right\|_1 = \left\|\left[\sqrt{\rho_{AB}}\left(\Gamma^A\otimes\Lambda_y^B\right)\sqrt{\rho_{AB}}\right]^T\right\|_1$$
$$= \left\|\sqrt{\rho_{AB}}\left((\Gamma^A)^T\otimes(\Lambda_y^B)^T\right)\sqrt{\rho_{AB}}\right\|_1. \quad (57)$$

One can easily prove for any Γ_A (not necessarily positive) that

$$\left(\sqrt{\rho_{AB}}\left((\Gamma^A)^T\otimes(\Lambda_y^B)^T\right)\sqrt{\rho_{AB}}\right)^R = \mathrm{Tr}_{AB}\left\{\left(\mathrm{id}\otimes\Gamma^A\otimes\Lambda_y^B\right)\Psi_{RAB}\right\}, \quad (58)$$

where Ψ_{RAB} is the canonical purification of ρ_{AB} defined as $\Psi_{RAB} \triangleq \sum_{x,x'}\sqrt{\lambda_x\lambda_{x'}}\,|x\rangle\langle x'|_{AB}\otimes|x\rangle\langle x'|_R$ for the spectral decomposition of ρ_{AB} given as $\rho_{AB} = \sum_x\lambda_x|x\rangle\langle x|_{AB}$ and $(\cdot)^R$ represents a state in the reference Hilbert space R. Now, using (58) we perform the following simplification

$$\sum_{y\in\mathcal{Y}}\left\|\sqrt{\rho_{AB}}\left(\Gamma^A\otimes\Lambda_y^B\right)\sqrt{\rho_{AB}}\right\|_1$$
$$\overset{(a)}{\leqslant} \sum_{y\in\mathcal{Y}^+}\left\|\sqrt{\rho_{AB}}\left(\Gamma^A\otimes\Lambda_y^B\right)\sqrt{\rho_{AB}}\right\|_1$$
$$= \sum_{y\in\mathcal{Y}^+}\left\|\mathrm{Tr}_{AB}\left\{\left(\mathrm{id}_R\otimes\Gamma^A\otimes\Lambda_y^B\right)\Psi_{RAB}\right\}\right\|_1$$
$$\overset{(b)}{=} \left\|\sum_{y\in\mathcal{Y}^+}\mathrm{Tr}_{AB}\left\{\left(\mathrm{id}_{RB}\otimes\Gamma^A\right)\left(\mathrm{id}_{RA}\otimes\Lambda_y^B\right)\Psi_{RAB}\otimes|y\rangle\langle y|\right\}\right\|_1$$
$$= \left\|\mathrm{Tr}_A\left\{\left(\mathrm{id}_{RY}\otimes\Gamma^A\right)\left(\sum_{y\in\mathcal{Y}^+}|y\rangle\langle y| \otimes\mathrm{Tr}_B\left\{\left(\mathrm{id}_{RA}\otimes\Lambda_y^B\right)\Psi_{RAB}\right\}\right)\right\}\right\|_1$$
$$\overset{(c)}{=} \left\|\mathrm{Tr}_A\left\{\left(\mathrm{id}_{RY}\otimes\Gamma^A\right)\sigma_{RAY}\right\}\right\|_1$$
$$\overset{(d)}{=} \left\|\mathrm{Tr}_{AZ}\left\{\left(\mathrm{id}_{RY}\otimes\Gamma^A\otimes\mathrm{id}_Z\right)\Phi_{RAYZ}^{\sigma_{RAY}}\right\}\right\|_1, \quad (59)$$

where (a) follows from the fact that $\left\|\sqrt{\rho_{AB}}\left(\Gamma^A\otimes\Lambda_{y_0}^B\right)\sqrt{\rho_{AB}}\right\|_1$ is always non-negative, (b) uses the triangle inequality for block diagonal operators, (c) uses σ_{RAY} defined as

$$\sigma_{RAY} = \sum_{y\in\mathcal{Y}^+}|y\rangle\langle y|\otimes\mathrm{Tr}_B\left\{\left(\mathrm{id}_{RA}\otimes\Lambda_y^B\right)\Psi_{RAB}\right\},$$

and finally, (d) uses $\Phi_{RAYZ}^{\sigma_{RAY}}$ defined as the canonical purification of σ_{RAY}. Note that the above inequality becomes an equality when $\sum_{y\in\mathcal{Y}}\Lambda_y = I$. Using similar sequence of arguments as used in (57) and (58), we have

$$\left\|\mathrm{Tr}_{AZ}\left\{\left(\mathrm{id}_{RY}\otimes\Gamma^A\otimes\mathrm{id}_Z\right)\Phi_{RAYZ}^{\sigma_{RAY}}\right\}\right\|_1$$
$$= \left\|\sqrt{\mathrm{Tr}_{RYZ}\{\Phi_{RAYZ}^{\sigma_{RAY}}\}}\,\Gamma^A\sqrt{\mathrm{Tr}_{RYZ}\{\Phi_{RAYZ}^{\sigma_{RAY}}\}}\right\|_1$$
$$= \|\sqrt{\rho_A}\,\Gamma^A\sqrt{\rho_A}\|_1.$$

This completes the proof.

B. Proof of Lemma 4

Let the operators of $\hat{M}_X$ and $\hat{M}_Y$ be denoted by $\{\hat{\Lambda}_i^X\}_{i\in\mathcal{I}}$ and $\{\hat{\Lambda}_j^Y\}_{j\in\mathcal{J}}$, respectively, and let the operators of M_X and M_Y be denoted by $\{\Lambda_i^X\}$ and $\{\Lambda_j^Y\}$, respectively, for some finite sets $\mathcal{I}$ and $\mathcal{J}$. With this notation, we need to show the following inequality

$$G \triangleq \sum_{i,j}\left\|\sqrt{\rho_{XY}}(\Lambda_i^X\otimes\Lambda_j^Y - \hat{\Lambda}_i^X\otimes\hat{\Lambda}_j^Y)\sqrt{\rho_{XY}}\right\|_1 + \mathrm{Tr}\left\{\Big(I-\sum_{i,j}\hat{\Lambda}_i^X\otimes\hat{\Lambda}_j^Y\Big)\rho_{XY}\right\} \leqslant (\epsilon_X+\epsilon_Y).$$

Next, by adding and subtracting appropriate terms, we get

$$\begin{aligned}
G &\leqslant \sum_{i,j}\left\|\sqrt{\rho_{XY}}(\Lambda_i^X\otimes\Lambda_j^Y - \hat{\Lambda}_i^X\otimes\Lambda_j^Y)\sqrt{\rho_{XY}}\right\|_1 \\
&\quad + \mathrm{Tr}\left\{\Big(I-\sum_i\hat{\Lambda}_i^X\Big)\rho_X\right\} \\
&\quad + \sum_{i,j}\left\|\sqrt{\rho_{XY}}(\hat{\Lambda}_i^X\otimes\Lambda_j^Y - \hat{\Lambda}_i^X\otimes\hat{\Lambda}_j^Y)\sqrt{\rho_{XY}}\right\|_1 \\
&\quad + \mathrm{Tr}\left\{\Big(I-\sum_j\hat{\Lambda}_j^Y\Big)\rho_Y\right\} + \mathrm{Tr}\left\{\Big(I-\sum_{i,j}\hat{\Lambda}_i^X\otimes\hat{\Lambda}_j^Y\Big)\rho_{XY}\right\} \\
&\quad - \mathrm{Tr}\left\{\Big(I-\sum_i\hat{\Lambda}_i^X\Big)\rho_X\right\} - \mathrm{Tr}\left\{\Big(I-\sum_j\hat{\Lambda}_j^Y\Big)\rho_Y\right\} \\
&\leqslant \sum_i\left\|\sqrt{\rho_X}(\Lambda_i^X-\hat{\Lambda}_i^X)\sqrt{\rho_X}\right\|_1 + \mathrm{Tr}\left\{\Big(I-\sum_i\hat{\Lambda}_i^X\Big)\rho_X\right\} \\
&\quad + \sum_j\left\|\sqrt{\rho_Y}(\Lambda_j^Y-\hat{\Lambda}_j^Y)\sqrt{\rho_Y}\right\|_1 + \mathrm{Tr}\left\{\Big(I-\sum_j\hat{\Lambda}_j^Y\Big)\rho_Y\right\} \\
&\quad + \mathrm{Tr}\left\{\Big(I-\sum_{i,j}\hat{\Lambda}_i^X\otimes\hat{\Lambda}_j^Y\Big)\rho_{XY}\right\} - \mathrm{Tr}\left\{\Big(I-\sum_i\hat{\Lambda}_i^X\Big)\rho_X\right\} \\
&\quad - \mathrm{Tr}\left\{\Big(I-\sum_j\hat{\Lambda}_j^Y\Big)\rho_Y\right\} \\
&\leqslant (\epsilon_X+\epsilon_Y) + \mathrm{Tr}\left\{\Big(\sum_i\hat{\Lambda}_i^X\otimes\big(I-\sum_j\hat{\Lambda}_j^Y\big)\Big)\rho_{XY}\right\} \\
&\qquad - \mathrm{Tr}\left\{\Big(I-\sum_j\hat{\Lambda}_j^Y\Big)\rho_Y\right\} \\
&\leqslant (\epsilon_X+\epsilon_Y),
\end{aligned}$$

where the second inequality follows by applying Lemma 3 twice, the third inequality follows from the hypotheses of the lemma, and the final inequality uses the fact that $\hat{M}^X$ and $\hat{M}^Y$ are sub-POVMs. This completes the proof of the lemma.

C. Proof of Lemma 7

Proof: Using the chain rule of quantum mutual information we see that

$$I(A;B|C,J)_\sigma = S(ACJ)_\sigma + S(BCJ)_\sigma - S(ABCJ)_\sigma - S(CJ)_\sigma.$$

The eigenvectors of the state σ_{ABCJ} are of the form $(0,\ldots,0,|j\rangle\otimes\left|v_i^j\right\rangle,0,\ldots,0)$, with eigenvalue $P_J(j)\lambda_i^j$, where $\left|v_i^j\right\rangle$ is an eigenvector of state ρ_{ABC}^j with eigenvalue λ_i^j. Hence

$$\begin{aligned}
S(ABCJ) &= -\sum_{j,i}P_J(j)\lambda_i^j\log(P_J(j)\lambda_i^j) \\
&\overset{(a)}{=} H(P_J) + \sum_{j=1}^n P_J(j)\sum_i[-\lambda_i^j\log\lambda_i^j] \\
&= S(J)_\sigma + \sum_{j=1}^n P_J(j)S(ABC)_{\rho^j},
\end{aligned}$$

where in (a) we used the grouping axiom of entropy. Applying similar arguments for $S(ACJ)$, $S(BCJ)$, and $S(CJ)$ we get the desired result. □

D. Proof of Lemma 8

Proof: Consider the trace norm expression given in (38). This expression can be bounded from above using the triangle inequality as

$$\begin{aligned}
&\left\|\sum_{w^n}\lambda_{w^n}\mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}}\frac{(1-\varepsilon)}{(1+\eta)}\sum_{l,\mu}\mathcal{T}_{W^{n,(\mu)}(l)}\right\|_1 \\
&\leqslant \left\|\sum_{w^n}\lambda_{w^n}\mathcal{T}_{w^n} - \frac{(1-\varepsilon)}{(1+\eta)}\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\tilde{P}_{W^n}(w^n)\mathcal{T}_{w^n}\right\|_1 \\
&\quad + \frac{(1-\varepsilon)}{(1+\eta)}\left\|\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\tilde{P}_{W^n}(w^n)\mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}}\sum_{l,\mu}\mathcal{T}_{W^{n,(\mu)}(l)}\right\|_1.
\end{aligned} \tag{60}$$

The first term in the right-hand side is bounded from above as

$$\begin{aligned}
&\left\|\sum_{w^n}\lambda_{w^n}\mathcal{T}_{w^n} - \frac{(1-\varepsilon)}{(1+\eta)}\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\tilde{P}_{W^n}(w^n)\mathcal{T}_{w^n}\right\|_1 \\
&\leqslant \left\|\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\lambda_{w^n}\left(1-\frac{1}{(1+\eta)}\right)\mathcal{T}_{w^n}\right\|_1 + \left\|\sum_{w^n\notin\mathcal{T}_\delta^{(n)}(W)}\lambda_{w^n}\mathcal{T}_{w^n}\right\|_1 \\
&\leqslant \left(\frac{\eta}{1+\eta}\right)\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\lambda_{w^n}\underbrace{\|\mathcal{T}_{w^n}\|_1}_{=1} + \sum_{w^n\notin\mathcal{T}_\delta^{(n)}(W)}\lambda_{w^n}\underbrace{\|\mathcal{T}_{w^n}\|}_{=1} \\
&\leqslant \left(\frac{\eta}{1+\eta}\right) + \varepsilon \leqslant \eta + \varepsilon \leqslant \frac{\epsilon}{2},
\end{aligned} \tag{61}$$

for all η sufficiently small and n sufficiently large. Now consider the second term in (60). Using the covering lemma from [39], this can be bounded as follows. For $w^n \in \mathcal{T}_\delta^{(n)}(W)$, let Π and Π_{w^n} denote the projectors onto the typical subspace of $\mathcal{T}^{\otimes n}$ and $\mathcal{T}_{w^n}$, respectively, where $\mathcal{T} = \sum_{w^n}\lambda_{w^n}\mathcal{T}_{w^n}$. From the definition of typical projectors, for any $\epsilon_1 \in (0,1)$ we have for sufficiently large n, the following inequalities satisfied for all $w^n \in \mathcal{T}_\delta^{(n)}(W)$:

$$\begin{aligned}
\mathrm{Tr}\,\Pi\mathcal{T}_{w^n} &\geqslant 1-\epsilon_1, \\
\mathrm{Tr}\,\Pi_{w^n}\mathcal{T}_{w^n} &\geqslant 1-\epsilon_1,
\end{aligned}$$

$$\begin{aligned}\operatorname{Tr}\Pi &\leqslant D,\\ \Pi_{w^n}\mathcal{T}_{w^n}\Pi_{w^n} &\leqslant \frac{1}{d}\Pi_{w^n},\end{aligned} \tag{62}$$

where $D = 2^{n(S(\mathcal{T})+\delta_1)}$ and $d = 2^{n\left[\left(\sum_w \lambda_w S(\mathcal{T}_w)\right)-\delta_2\right]}$, and $\delta_1(\delta) \searrow 0, \delta_2(\delta) \searrow 0$ as $\delta \searrow 0$. From the statement of the covering lemma, we know that for an ensemble $\{\tilde{P}_{W^n}(w^n), \mathcal{T}_{w^n}\}_{w^n\in\mathcal{W}^n}$, if there exists projectors Π and Π_{w^n} such that they satisfy the set of inequalities in (62), then for all sufficiently large n, if $n(R+C) > \log_2 \frac{D}{d}$, the obfuscation error, defined as

$$\left\| \sum_{w^n} \tilde{P}_W^n(w^n)\mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}}\sum_{l,\mu}\mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1,$$

can be made smaller than $\epsilon_1 + 4\sqrt{\epsilon_1} + 24\sqrt[4]{\epsilon_1}$ with high probability. This gives us the following rate constraints $R + C > S(\sum_w \lambda_w \mathcal{T}_w) - \sum_w \lambda_w S(\mathcal{T}_w) + \delta_1 + \delta_2 = \chi\left(\{\lambda_w\}, \{\hat{\rho}_w \otimes \sigma_w\}\right) + \delta_1 + \delta_2$. Using this constraint and the bound from (61), the result follows. □

Appendix C
Proof of Propositions

A. Proof of Proposition 1

The second term in the trace distance in S_2 can be expressed as

$$\begin{aligned}
&(\mathrm{id}\otimes[\tilde{M}_{AB}])(\Psi^\rho_{RAB})\\
&= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i,j}\Phi_{F^{(\mu_1,\mu_2)}(i,j)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes\Gamma_i^{A,(\mu_1)}\otimes\Gamma_j^{B,(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad + (1-\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\{(\mathrm{id}\otimes\mathrm{id}\otimes\mathrm{id})\Psi^\rho_{RAB}\}\\
&= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\\
&\quad\times\Bigg[\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i,j\geqslant 1}\sum_{(u,v)\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j)}\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{j\geqslant 1}\sum_{v\in\mathcal{B}_2^{(\mu_2)}(j)}\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i\geqslant 1}\sum_{u\in\mathcal{B}_1^{(\mu_1)}(i)}\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi^\rho_{RAB}\right\}\Bigg]\\
&\quad + (1-\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\{(\mathrm{id}\otimes\mathrm{id}\otimes\mathrm{id})\Psi^\rho_{RAB}\}.
\end{aligned} \tag{63}$$

Similarly, for the first term within the trace distance in S_2, we have

$$\begin{aligned}
&\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}(\mathrm{id}\otimes[M_1^{(\mu_1)}]\otimes[M_2^{(\mu_2)}])(\Psi^\rho_{RAB})\\
&= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\\
&\quad\times\Bigg[\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Phi_{(u,v)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{v\in\mathcal{V}}\Phi_{(0_U,v)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\Phi_{(u,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi^\rho_{RAB}\right\}\\
&\quad+\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi^\rho_{RAB}\right\}\Bigg]\\
&\quad+(1-\mathbb{1}_{\{\text{sP}-1\}})\mathbb{1}_{\{\text{sP}-2\}}\frac{1}{N_2}\sum_{\mu_2}\sum_{v\in\mathcal{V}\cup\{0_V\}}\Phi_{(0_U,v)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\{\mathrm{id}\otimes\mathrm{id}\otimes B_v^{(\mu_2)})\Psi_{RAB}\}\\
&\quad+(1-\mathbb{1}_{\{\text{sP}-2\}})\mathbb{1}_{\{\text{sP}-1\}}\frac{1}{N_1}\sum_{\mu_1}\sum_{u\in\mathcal{U}\cup\{0_U\}}\Phi_{(u,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\{\mathrm{id}\otimes A_u^{(\mu_1)}\otimes\mathrm{id})\Psi_{RAB}\}\\
&\quad+(1-\mathbb{1}_{\{\text{sP}-2\}})(1-\mathbb{1}_{\{\text{sP}-1\}})\Phi_{(0_U,0_V)}\\
&\qquad\otimes \operatorname{Tr}_{AB}\{(\mathrm{id}\otimes\mathrm{id}\otimes\mathrm{id})\Psi_{RAB}\}.
\end{aligned} \tag{64}$$

By replacing the terms in S_2 using the corresponding expansions from (63) and (64), we observe that the fourth terms on the right hand side of (63) get canceled with the corresponding terms on the right hand side of (64). Next we take the second term in (63) and apply the triangle inequality and bound from above its l_1 norm by

$$\frac{1}{N_1}\sum_{\mu_1}\operatorname{Tr}\left(\left(I-\sum_{u\in\mathcal{U}}A_u^{(\mu_1)}\right)\rho_A\right).$$

Similarly, we can bound the rest of the terms in (63), (64), except the first terms. The l_1 norm of the difference of the first terms in (63), (64) can be written as

$$\begin{aligned}
&\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Big\|(\Phi_{(u,v)}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)})\\
&\qquad\otimes \operatorname{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^\rho_{RAB}\right\}\Big\|_1\\
&= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{(u,v)}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1
\end{aligned}$$

$$\times \operatorname{Tr}_{RAB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi^{\rho}_{RAB}\right\}$$
$$= \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2} \times \sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},$$

where the first equality is obtained by using the definition of trace norm and the last equality follows from the definition of $A_u^{(\mu_1)}$ and $B_v^{(\mu_2)}$, with $\Omega_{u,v}$ as given in the statement of the proposition. This completes the proof.

B. Proof of Proposition 2

Using the proof of Theorem 7, one can show that

$$\frac{2}{N_1}\sum_{\mu_1}\operatorname{Tr}\left((I-\sum_{u\in\mathcal{U}}A_u^{(\mu_1)})\rho_A\right)+\frac{2}{N_2}\sum_{\mu_2}\operatorname{Tr}\left((I-\sum_{v\in\mathcal{V}}B_v^{(\mu_2)})\rho_B\right) + 2(2-\mathbb{1}_{\{\text{sP-1}\}}-\mathbb{1}_{\{\text{sP-2}\}})\leqslant \alpha_A+\alpha_B.$$

Recall from Proposition 1 that S_3 can be simplified as

$$S_3 = \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2} \times \sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1\gamma_{u^n}^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},$$

For any (u,v), the 1-norm above can be bounded from above by the following quantity:

$$\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\|_1\leqslant 2[\mathbb{1}_{\{(u,v)\notin\mathcal{W}\}}+\mathbb{1}^{(\mu_1,\mu_2)}(u,v)],$$

where $\mathbb{1}^{(\mu_1,\mu_2)}(u,v)\triangleq$

$$\mathbb{1}\left\{\exists(\tilde{u},\tilde{v},i,j):(u,v)\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j),(\tilde{u},\tilde{v})\in\mathcal{C}^{(\mu_1,\mu_2)}\bigcap\mathcal{W}, (\tilde{u},\tilde{v})\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j),(\tilde{u},\tilde{v})\neq(u,v)\right\}.$$

Using such indicator functions, S_3 can be bounded from above as $S_3\leqslant S_4+S_5$, where

$$S_4\triangleq \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{2}{N_1N_2}\sum_{(u,v)}\Omega_{u,v}\sum_{\mu_1,\mu_2}\mathbb{1}_{\{(u,v)\notin\mathcal{W}\}}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)},$$

$$S_5\triangleq\frac{(1-\theta_1)(1-\theta_2)}{(1+\epsilon_1)(1+\epsilon_2)K_1K_2}\sum_{l,k}\sum_{(u,v)}\Omega_{u,v}\frac{2}{N_1N_2} \times\sum_{\mu_1,\mu_2}\mathbb{1}^{(\mu_1,\mu_2)}(u,v)\mathbb{1}\{U^{(\mu_1)}(l)=u,V^{(\mu_2)}(k)=v\},$$

where we have bounded the indicator random variables in S_5. We provide bounds on the expectation of S_4 and S_5. For that we take the expectation of the indicator functions with respect to random variables which are independent of each other and distributed according to $\{\lambda_u^A\}_{u\in\mathcal{U}}$, and $\{\lambda_v^B\}_{v\in\mathcal{V}}$. First consider the following argument:

$$S_4\leqslant\left|S_4-2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}\right|+2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}$$
$$\overset{(a)}{\leqslant}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left|\lambda_{u,v}^{AB}-\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\right|$$
$$+\mathbb{1}_{\{\text{sP}-1\}}\mathbb{1}_{\{\text{sP}-2\}}(1-\frac{1}{N_1N_2}\sum_{u,v}\sum_{\mu_1,\mu_2}\gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}) +(1-\mathbb{1}_{\{\text{sP}-1\}}\mathbb{1}_{\{\text{sP}-2\}})+2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}$$
$$\overset{(b)}{\leqslant} S_1+2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB},$$

where (a) follows from the two different definitions of variational distance between probability distributions, (b) follows from Lemma 5. Taking expectation we obtain

$$\mathbb{E}[S_4]\leqslant(\alpha_A+\alpha_B)+2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}, \tag{65}$$

where we used the bound developed earlier on S_1 in the mutual covering lemma. For S_5 we have

$$\mathbb{E}\Big[\mathbb{1}^{(\mu_1,\mu_2)}(u,v)\mathbb{1}\left\{U^{(\mu_1)}(l)=u,V^{(\mu_2)}(k)=v\right\}\Big]$$
$$\overset{(a)}{\leqslant}\sum_{\substack{(\tilde{u},\tilde{v})\in\mathcal{W}\\(\tilde{u},\tilde{v})\neq(u,v)}}\sum_{i,j}\sum_{(\tilde{l},\tilde{k})}\mathbb{E}\Big[\mathbb{1}\left\{(u,v)\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j)\right\}$$
$$\times\mathbb{1}\left\{(\tilde{u},\tilde{v})\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j)\right\}$$
$$\times\mathbb{1}\left\{U^{(\mu_1)}(l)=u,V^{(\mu_2)}(k)=v\right\}$$
$$\times\mathbb{1}\left\{U^{(\mu_1)}(\tilde{l})=\tilde{u},V^{(\mu_2)}(\tilde{k})=\tilde{v}\right\}\Big]$$
$$\overset{(b)}{\leqslant}\frac{\lambda_u^A\lambda_v^B}{(1-\theta_1)(1-\theta_2)}\Bigg[\frac{\lambda_{\mathrm{m}}^A\lambda_{\mathrm{m}}^B|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2}+\frac{K_1W_A\lambda_{\mathrm{m}}^A}{(1-\theta_1)T_1} \times\left(1+\frac{\lambda_{\mathrm{m}}^BK_2}{(1-\theta_2)}\right)+\frac{K_2W_B\lambda_{\mathrm{m}}^B}{(1-\theta_2)T_2}\left(1+\frac{\lambda_{\mathrm{m}}^AK_1}{(1-\theta_1)}\right)\Bigg], \tag{66}$$

where (a) follows from the union bound, and (b) follows by noting that there are 5 cases to consider, and by evaluating the expectation of the indicator functions while recalling $W_A=\max_{v\in\mathcal{V}}|\{u:(u,v)\in\mathcal{W}\}|$, and $W_B=\max_{u\in\mathcal{U}}|\{v:(u,v)\in\mathcal{W}\}|$, $\lambda_{\mathrm{m}}^A=\max_u\lambda_u^A$, $\lambda_{\mathrm{m}}^B=\max_v\lambda_v^B$. This implies that

$$\mathbb{E}[S_5]\leqslant\frac{2}{(1+\epsilon_1)(1+\epsilon_2)}\Bigg[\frac{\lambda_{\mathrm{m}}^A\lambda_{\mathrm{m}}^B|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2} +\frac{K_1W_A\lambda_{\mathrm{m}}^A}{(1-\theta_1)T_1}\left(1+\frac{\lambda_{\mathrm{m}}^BK_2}{(1-\theta_2)}\right)+\frac{K_2W_B\lambda_{\mathrm{m}}^B}{(1-\theta_2)T_2} \times\left(1+\frac{\lambda_{\mathrm{m}}^AK_1}{(1-\theta_1)}\right)\Bigg]\times\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Omega_{u,v}\lambda_u^A\lambda_v^B.$$

We have the following lemma.

Lemma 11: We have

$$\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Omega_{u,v}\lambda_u^A\lambda_v^B\leqslant\frac{f_1f_2}{F_1F_2}.$$

Proof: Firstly, note that

$$\sum_{\substack{u\in\mathcal{U}\\v\in\mathcal{V}}}\Omega_{u,v}\lambda_u^A\lambda_v^B=\operatorname{Tr}\Bigg\{\Bigg[\sqrt{\rho_A}^{-1}\Big(\sum_{u\in\mathcal{U}}\lambda_u^A\tilde{\rho}_u^A\Big)\sqrt{\rho_A}^{-1} \otimes\sqrt{\rho_B}^{-1}\Big(\sum_{v\in\mathcal{V}}\lambda_v^B\tilde{\rho}_v^B\Big)\sqrt{\rho_B}^{-1}\Bigg]\rho_{AB}\Bigg\}. \tag{67}$$

Consider,

$$\sum_{u\in\mathcal{U}}\lambda_u^A\tilde{\rho}_u^A = \hat{\Pi}^A\Pi_{\rho_A}\left(\sum_{u\in\mathcal{U}}\lambda_u^A\Pi_u^A\hat{\rho}_u^A\Pi_u^A\right)\Pi_{\rho_A}\hat{\Pi}^A \overset{(a)}{\leqslant} \hat{\Pi}^A\Pi_{\rho_A}\left(\sum_u\lambda_u^A\hat{\rho}_u^A\right)\Pi_{\rho_A}\hat{\Pi}^A \overset{(b)}{\leqslant} \hat{\Pi}^A\Pi_{\rho_A}\rho_A\Pi_{\rho_A}\hat{\Pi}^A \overset{(c)}{\leqslant} \frac{1}{F_1}\hat{\Pi}^A\Pi_{\rho_A}\hat{\Pi}^A \overset{(d)}{\leqslant} \frac{1}{F_1}\Pi_{\rho_A}.$$

where (a) follows from the hypothesis $\Pi_u^A\hat{\rho}_u^A\Pi_u^A \leqslant \hat{\rho}_u^A$, (b) from the fact that M_A is a sub-POVM, and (c) from the hypothesis $\Pi_{\rho_A}\rho_A\Pi_{\rho_A} \leqslant \frac{1}{F_1}\Pi_{\rho_A}$, and (d) from the commutativity of $\hat{\Pi}^A$ and Π_{ρ_A}, where the commutativity follows from the fact that $\hat{\Pi}^A$ is a cut-off projector on the subspace determined by Π_{ρ_A}. This implies that

$$\sqrt{\rho_A}^{-1}\left(\sum_{u\in\mathcal{U}}\lambda_{u^n}^A\tilde{\rho}_u^A\right)\sqrt{\rho_A}^{-1} \leqslant \frac{1}{F_1}\sqrt{\rho_A}^{-1}\Pi_{\rho_A}\sqrt{\rho_A}^{-1} \leqslant \frac{f_1}{F_1}\Pi_{\rho_A}, \tag{68}$$

where the last inequality follows by using the hypothesis $\sqrt{\rho_A}^{-1}\Pi_{\rho_A}\sqrt{\rho_A}^{-1} \leqslant f_1\Pi_{\rho_A}$. Using the same arguments for the operators acting on $\mathcal{H}_B$, we have

$$\sqrt{\rho_B}^{-1}\left(\sum_{v\in\mathcal{V}}\lambda_{v^n}^B\tilde{\rho}_v^B\right)\sqrt{\rho_B}^{-1} \leqslant \frac{1}{F_2}\sqrt{\rho_B}^{-1}\Pi_{\rho_B}\sqrt{\rho_B}^{-1} \leqslant \frac{f_2}{F_2}\Pi_{\rho_B}. \tag{69}$$

Using (68) and (69) in (67), gives

$$\sum_{u,v}\Omega_{u,v}\lambda_u^A\lambda_v^B \leqslant \frac{f_1f_2}{F_1F_2}\operatorname{Tr}\left(\Pi_{\rho_A}\otimes\Pi_{\rho_B}\right)\rho_{AB} \leqslant \frac{f_1f_2}{F_1F_2}\operatorname{Tr}\rho_{AB} = \frac{f_1f_2}{F_1F_2},$$

which is the desired result. □

C. Proof of Proposition 4

Fix an arbitrary $\epsilon > 0$, and $\eta,\delta \in (0,1)$ sufficiently small. Recalling $S_2(\tilde{\mu}_1,\tilde{\mu}_2)$, we have

$$\begin{aligned}
&S_2(\tilde{\mu}_1,\tilde{\mu_2}) \\
&\leqslant \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{z^n}\sum_{u^n,v^n}\left|P^n_{Z|U,V}(z^n|u^n,v^n) - P^n_{Z|U,V}\left(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\right)\right| \\
&\qquad\times\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 \\
&\leqslant \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n} \\
&\qquad\times\sum_{z^n}\left|P^n_{Z|U,V}(z^n|u^n,v^n) - P^n_{Z|U,V}\left(z^n|e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\right)\right| \\
&\leqslant \frac{2}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\Big(\mathbb{1}_{\{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)\}} + \mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\Big) \\
&\qquad\times\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n},
\end{aligned} \tag{70}$$

where Ω_{u^n,v^n} and $\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)$ are defined as

$$\Omega_{u^n,v^n} \triangleq \operatorname{Tr}\left\{\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}(\tilde{\rho}_{u^n}^A\otimes\tilde{\rho}_{v^n}^B)\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}\rho_{AB}^{\otimes n}\right\},$$

$$\begin{aligned}
&\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n) \\
&\triangleq \mathbb{1}\Big\{\exists(\tilde{u}^n,\tilde{v}^n,i,j):(u^n,v^n)\in\mathcal{B}_1^{(\bar{\mu}_1)}(i)\times\mathcal{B}_2^{(\bar{\mu}_2)}(j),(\tilde{u}^n,\tilde{v}^n)\in \\
&\quad\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}\bigcap\mathcal{T}_\delta^{(n)}(UV),(\tilde{u}^n,\tilde{v}^n)\in\mathcal{B}_1^{(\bar{\mu}_1)}(i)\times\mathcal{B}_2^{(\bar{\mu}_2)}(j), \\
&\quad(\tilde{u}^n,\tilde{v}^n)\neq(u^n,v^n)\Big\}.
\end{aligned}$$

Now we can use the bound $S_2 \leqslant S_{21} + S_{22}$, where

$$\begin{aligned}
&S_{21}(\tilde{\mu}_1,\tilde{\mu}_2) \\
&\triangleq \frac{2}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\mathbb{1}_{\{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)\}}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}, \\
&S_{22}(\tilde{\mu}_1,\tilde{\mu}_2) \\
&\triangleq \frac{2}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u^n,v^n}\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}.
\end{aligned}$$

We begin by bounding the term corresponding to S_{21}. Consider the following argument.

$$\begin{aligned}
&\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}S_{21}\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} \\
&\leqslant \left|\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}S_{21}\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} - \sum_{\substack{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(UV)\\ u^n\in\mathcal{T}_\delta^{(n)}(U),v^n\in\mathcal{T}_\delta^{(n)}(V)}}2\lambda_{u^n,v^n}^{AB}\right| \\
&\qquad+\sum_{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)}2\lambda_{u^n,v^n}^{AB} \\
&\overset{(a)}{\leqslant} 2\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\Bigg|\lambda_{u^n,v^n}^{AB} - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} \\
&\qquad\times\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}\Bigg| + \sum_{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(UV)}2\lambda_{u,v}^{AB} \\
&\overset{(b)}{\leqslant} 2\tilde{S}_1 + 2\sum_{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(UV)}\lambda_{u^n,v^n}^{AB},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{S}_1 \triangleq \Bigg\|&(\mathrm{id}\otimes M_A^{\otimes n}\otimes M_B^{\otimes n})(\Psi_{RAB}^\rho)^{\otimes n} \\
&-\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}(\mathrm{id}\otimes[M_1^{(\bar{\mu}_1)}]\otimes[M_2^{(\bar{\mu}_2)}])(\Psi_{RAB}^\rho)^{\otimes n}\Bigg\|_1,
\end{aligned}$$

(a) follows by applying the triangle inequality, and (b) follows from Lemma 5. Note that in $\tilde{S}_1$, the average over the entire common information sequence $(\bar{\mu}_1,\bar{\mu}_2)$ is inside the norm. Using the Lemmas 2 and 4, and the proof of Theorem 7, for any $\epsilon \in (0,1)$, and any $\eta,\delta \in (0,1)$ sufficiently small, and any n sufficiently large, if

$$\begin{aligned}
&\tilde{R}_1 > I(U;RB)_{\sigma_1},\quad \tilde{R}_2 > I(V;RA)_{\sigma_2}, \\
&\tilde{R}_1 + \frac{1}{n}\log(\bar{N}_1) > S(U)_{\sigma_3},\quad \tilde{R}_2 + \frac{1}{n}\log(\bar{N}_2) > S(V)_{\sigma_3},
\end{aligned} \tag{71}$$

then $\mathbb{E}[\tilde{S}_1] \leqslant \epsilon$. Consequently, we have

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}\mathbb{E}\left[S_{21}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\right]\leqslant 4\epsilon.$$

Now considering the term S_{22}, using a simplification similar to (66) we obtain

$$\mathbb{E}\left[\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\mathbb{1}_{\{U^{n,(\mu_1)}(l)=u^n\}}\mathbb{1}_{\{V^{n,(\mu_2)}(k)=v^n\}}\right] \leqslant \frac{5\ \lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)^2(1-\varepsilon')^2}2^{-n(I(U;V)-3\delta_1)}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}.$$

Substituting this in the expression for S_{22} gives

$$\begin{aligned}\mathbb{E}[S_{22}] &\leqslant 10\frac{2^{-n(I(U;V)-3\delta_1)}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}}{(1+\eta)^2(1-\varepsilon)^2(1-\varepsilon')^2} \times \sum_{u^n,v^n}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n}\\ &\leqslant 10\frac{2^{-n(I(U;V)-3\delta_1-\delta_{AB})}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}}{(1+\eta)^2(1-\varepsilon)^2(1-\varepsilon')^2},\end{aligned}$$

where the second inequality above uses arguments similar to Lemma 11. Therefore, if

$$\tilde{R}_1+\tilde{R}_2-R_1-R_2 \leqslant I(U;V)_{\sigma_3}-3\delta_1-\delta_{AB}-\delta, \quad (72)$$

then we have $\mathbb{E}[S_{22}] \leqslant 10\frac{2^{-n\delta}}{(1+\eta)^2(1-\varepsilon)(1-\varepsilon')} < \epsilon$, for all sufficiently large n. Hence

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}\mathbb{E}(S_2(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)) < 5\epsilon,$$

for all sufficiently large n, if (71) and (72) are satisfied.

D. Proof of Proposition 5

We bound $\widetilde{S}$ as $\widetilde{S} \leqslant \widetilde{S}_2+\widetilde{S}_3+\widetilde{S}_4$, where

$$\widetilde{S}_2 \triangleq \left\|\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{i>0}\sqrt{\rho^{\otimes n}_{AB}}\left(\Gamma^{A,(\mu_1)}_i\otimes\Gamma^{B,(\mu_2)}_0\right) \times\sqrt{\rho^{\otimes n}_{AB}}P^n_{Z|U,V}(z^n|u^n_0,v^n_0)\right\|_1,$$

$$\widetilde{S}_3 \triangleq \left\|\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{j>0}\sqrt{\rho^{\otimes n}_{AB}}\left(\Gamma^{A,(\mu_1)}_0\otimes\Gamma^{B,(\mu_2)}_j\right) \times\sqrt{\rho^{\otimes n}_{AB}}P^n_{Z|U,V}(z^n|u^n_0,v^n_0)\right\|_1,$$

$$\widetilde{S}_4 \triangleq \left\|\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho^{\otimes n}_{AB}}\left(\Gamma^{A,(\mu_1)}_0\otimes\Gamma^{B,(\mu_2)}_0\right) \times\sqrt{\rho^{\otimes n}_{AB}}P^n_{Z|U,V}(z^n|u^n_0,v^n_0)\right\|_1.$$

Analysis of $\widetilde{S}_2$: We have

$$\begin{aligned}&\widetilde{S}_2\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\\ &\leqslant \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{z^n}P^n_{Z|U,V}(z^n|u^n_0,v^n_0) \times\left\|\sqrt{\rho^{\otimes n}_{AB}}\left(\sum_{i>0}\Gamma^{A,(\mu_1)}_i\otimes\Gamma^{B,(\mu_2)}_0\right)\sqrt{\rho^{\otimes n}_{AB}}\right\|_1\\
&\overset{(a)}{=} \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2} \times\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho^{\otimes n}_{AB}}\left(\sum_{u^n}A^{(\mu_1)}_{u^n}\otimes\Gamma^{B,(\mu_2)}_0\right)\sqrt{\rho^{\otimes n}_{AB}}\right\|_1\\
&\leqslant \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2} \times\sum_{\mu_1,\mu_2}\sum_{u^n}\left\|\sqrt{\rho^{\otimes n}_{AB}}\left(A^{(\mu_1)}_{u^n}\otimes\Gamma^{B,(\mu_2)}_0\right)\sqrt{\rho^{\otimes n}_{AB}}\right\|_1\\
&\overset{(b)}{\leqslant} \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho^{\otimes n}_{B}}\Gamma^{B,(\mu_2)}_0\sqrt{\rho^{\otimes n}_{B}}\right\|_1\\
&= \frac{1}{N_2}\sum_{\mu_2}\left\|\sum_{v^n}\lambda^B_{v^n}\hat{\rho}^B_{v^n}-\sum_{v^n}\sqrt{\rho^{\otimes n}_{B}}B^{(\mu_2)}_{v^n}\sqrt{\rho^{\otimes n}_{B}}\right\|_1\\
&\overset{(c)}{\leqslant} \frac{1}{N_2}\sum_{\mu_2}\left\|\sum_{v^n}\lambda^B_{v^n}\hat{\rho}^B_{v^n}-\frac{(1-\varepsilon')}{(1+\eta)}\frac{1}{2^{n\tilde{R}_2}}\sum_{k=1}^{2^{n\tilde{R}_2}}\hat{\rho}^B_{V^{n,(\mu_2)}(k)}\right\|_1 + \underbrace{\frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\zeta^{(\mu_2)}_{v^n}\left\|\hat{\rho}^B_{v^n}-\tilde{\rho}^B_{v^n}\right\|_1}_{\tilde{S}_{22}}, \quad (73)\end{aligned}$$

where (a) uses the fact that $\sum_{i>0}\Gamma^{A,(\mu_1)}_i = \sum_{u^n}A^{(\mu_1)}_{u^n}$, (b) uses the fact that under the event $\{\mathbb{1}_{\{\text{sP-1}\}}=1\}$, we have $\sum_{u^n}A^{(\mu_1)}_{u^n} \leqslant I$, and Lemma 3. Finally (c) follows from adding and subtracting an appropriate term. Regarding the first term in (73) using Lemma 8 we claim that for for any $\epsilon \in (0,1)$, any $\eta,\delta \in (0,1)$ sufficiently small, and any n sufficiently large, the term can be made smaller than ϵ, if $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, where σ_2 is as defined in the statement of the theorem. Note that the requirement we obtain on $\tilde{R}_2$ here was already imposed earlier in Proposition 3. And as for the second term, we use the gentle measurement lemma and bound its expected value as

$$\mathbb{E}\left[\frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\zeta^{(\mu_2)}_{v^n}\left\|\hat{\rho}^B_{v^n}-\tilde{\rho}^B_{v^n}\right\|_1\right] = \sum_{v^n\in\mathcal{T}^{(n)}_\delta(V)}\frac{\lambda^B_{v^n}}{(1+\eta)}\left\|\hat{\rho}^B_{v^n}-\tilde{\rho}^B_{v^n}\right\|_1 \leqslant \epsilon_{\tilde{S}_2},$$

where the inequality is based on the repeated usage of the average gentle measurement lemma by setting $\epsilon_{\tilde{S}_2} = \frac{(1-\varepsilon')}{(1+\eta)}(2\sqrt{\varepsilon'_B}+2\sqrt{\varepsilon''_B})$ with $\epsilon_{\tilde{S}_2} \searrow 0$ as $n\to\infty$ and $\varepsilon'_B = \varepsilon'_p+2\sqrt{\varepsilon'_p}$ and $\varepsilon''_B = 2\varepsilon'_p+2\sqrt{\varepsilon'_p}$ for $\varepsilon'_p \triangleq 1-\min\left\{\mathrm{Tr}\,\Pi_{\rho_B}\hat{\rho}^B_{v^n},\mathrm{Tr}\,\Pi_{v^n}\hat{\rho}^B_{v^n},1-\varepsilon'\right\}$. Hence $\mathbb{E}\left[\tilde{S}_2\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right] \leqslant 2\epsilon$.

Analysis of $\widetilde{S}_3$: Due to the symmetry in $\widetilde{S}_2$ and $\widetilde{S}_3$, the analysis of $\widetilde{S}_3$ follows very similar arguments as that of $\widetilde{S}_2$ and hence we skip it.

Analysis of $\widetilde{S}_4$: We have

$$\begin{aligned}&\widetilde{S}_4\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\\ &\leqslant \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{z^n}P^n_{Z|U,V}(z^n|u^n_0,v^n_0) \times\left\|\sqrt{\rho^{\otimes n}_{AB}}\left(\Gamma^{A,(\mu_1)}_0\otimes\Gamma^{B,(\mu_2)}_0\right)\sqrt{\rho^{\otimes n}_{AB}}\right\|_1\end{aligned}$$

$$\leqslant \frac{\mathbb{1}_{\{\text{sP-1}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes I\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 + \frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1, \quad (74)$$

where the inequalities above are obtained by a straight forward substitution and use of triangle inequality. With the above constraints on $\tilde{R}_1$ and $\tilde{R}_2$, we have $0 \leqslant \Gamma_0^{A,(\mu_1)} \leqslant I$ and $0 \leqslant \Gamma_0^{B,(\mu_2)} \leqslant I$. This simplifies the first term in (74) as

$$\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes I\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 = \frac{1}{N_1}\sum_{\mu_1}\left\|\sqrt{\rho_{A}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\right)\sqrt{\rho_{A}^{\otimes n}}\right\|_1.$$

Similarly, the second term in (74) simplifies using Lemma 3 as

$$\frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 \leqslant \frac{1}{N_1}\sum_{\mu_1}\left\|\sqrt{\rho_{A}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\right)\sqrt{\rho_{A}^{\otimes n}}\right\|_1.$$

Using these simplifications, we have

$$\widetilde{S}_4\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} \leqslant \frac{2}{N_1}\sum_{\mu_1}\left\|\sqrt{\rho_{A}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\right)\sqrt{\rho_{A}^{\otimes n}}\right\|_1.$$

The above expression is similar to the one obtained in the simplification of $\widetilde{S}_2$ and hence we can bound $\widetilde{S}_4$ using the same constraints as $\widetilde{S}_2$.

E. Proof of Proposition 6

Note that from triangle inequality, we have $Q_1 \leqslant J$. Further, we add and subtract an appropriate term within J and use triangle inequality obtain $J \leqslant J_1 + J_2$, where J_1 and J_2 are shown at the bottom of the page. Now with the intention of employing Lemma 8, we express J_1 as

$$J_1 = \Bigg\| \sum_{z^n,u^n,v^n}\lambda_{u^n,v^n}^{AB}\hat{\rho}_{u^n,v^n}^{AB}\otimes P_{Z|U,V}^n(z^n|u^n,v^n)\,|v^n\rangle\langle v^n|\otimes|z^n\rangle\langle z^n| - \frac{(1-\varepsilon)}{(1+\eta)}\frac{1}{2^{n(\tilde{R}_1+C_1)}}\sum_{\mu_1,l}\sum_{z^n,u^n,v^n}\mathbb{1}_{\{U^{n,(\mu_1)}(l)=u^n\}}\frac{\lambda_{u^n,v^n}^{AB}}{\lambda_{u^n}^{A}} \times \hat{\rho}_{u^n,v^n}^{AB}\otimes P_{Z|U,V}^n(z^n|u^n,v^n)\,|v^n\rangle\langle v^n|\otimes|z^n\rangle\langle z^n| \Bigg\|_1,$$

where the equality above is obtained by using the definitions of $\gamma_{u^n}^{(\mu_1)}$ and $\hat{\rho}_{u^n,v^n}^{AB}$, followed by using the triangle inequality for the block diagonal operators, which in fact becomes an equality. Let us define $\mathcal{T}_{u^n}$ as

$$\mathcal{T}_{u^n} \triangleq \sum_{z^n,v^n}\frac{\lambda_{u^n,v^n}^{AB}}{\lambda_{u^n}^{A}}\hat{\rho}_{u^n,v^n}^{AB}\otimes P_{Z|U,V}^n(z^n|u^n,v^n)\,|v^n\rangle\langle v^n|\otimes|z^n\rangle\langle z^n|.$$

Note that the above definition of $\mathcal{T}_{u^n}$ contains all the elements in product form, and thus it can be written as $\mathcal{T}_{u^n} = \bigotimes_{i=1}^n \mathcal{T}_{u_i}$. This simplifies J_1 as

$$J_1 = \left\|\sum_{u^n}\lambda_{u^n}^{A}\mathcal{T}_{u^n} - \frac{(1-\varepsilon)}{(1+\eta)}\frac{1}{2^{n(\tilde{R}_1+C_1)}}\sum_{\mu_1,l}\mathcal{T}_{U^{n,(\mu_1)}(l)}\right\|_1.$$

Now, using Lemma 8 we get the following bound. For any $\epsilon \in (0,1)$, any $\eta,\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}(J_1) < \epsilon$ if

$$\tilde{R}_1 + C_1 > S\left(\sum_{u\in\mathcal{U}}\lambda_u^A\mathcal{T}_u\right) + \sum_{u\in\mathcal{U}}\lambda_u^A S(\mathcal{T}_u) = I(U;RZV)_{\sigma_3}, \quad (75)$$

where $\sigma_3 = \sum_{u\in\mathcal{U}}\lambda_u^A\mathcal{T}_u\otimes|u\rangle\langle u|$.

Now, we consider the term corresponding to J_2 and prove that its expectation with respect to the Alice's codebook is small. Recalling J_2, we get

$$J_2 \leqslant \frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\sum_{u^n,v^n}\sum_{z^n}P_{Z|U,V}^n(z^n|u^n,v^n) \times \left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^{A}}\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B - A_{u^n}^{(\mu_1)}\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1$$
$$= \frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\sum_{u^n,v^n}\gamma_{u^n}^{(\mu_1)}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\left(\frac{1}{\lambda_{u^n}^{A}}\Lambda_{u^n}^A - \sqrt{\rho_{A}^{\otimes n}}^{-1}\tilde{\rho}_{u^n}^A\sqrt{\rho_{A}^{\otimes n}}^{-1}\right)\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1,$$

where the inequality is obtained by using triangle and the next equality follows from the fact that $\sum_{z^n}P_{Z|U,V}^n(z^n|u^n,v^n) = 1$ for all $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ and using the definition of $A_{u^n}^{(\mu_1)}$. By applying expectation of J_2 over the Alice's codebook, we get

$$\mathbb{E}[J_2] \leqslant \frac{1}{(1+\eta)}\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}\lambda_{u^n}^A\sum_{v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\left(\frac{1}{\lambda_{u^n}^{A}}\Lambda_{u^n}^A - \sqrt{\rho_{A}^{\otimes n}}^{-1}\tilde{\rho}_{u^n}^A\sqrt{\rho_{A}^{\otimes n}}^{-1}\right)\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1,$$

$$J_1 \triangleq \sum_{z^n,v^n}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B - \frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^{A}}\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1,$$

$$J_2 \triangleq \sum_{z^n,v^n}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^{A}}\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B - \frac{1}{N_1}\sum_{\mu_1=1}^{N_1}A_{u^n}^{(\mu_1)}\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1.$$

where we have used the fact that $\mathbb{E}[\gamma_{u^n}^{(\mu_1)}] = \frac{\lambda_{u^n}^A}{(1+\eta)}$. To simplify the above equation, we employ Lemma 3 from Section IV-B.2 that completely discards the effect of Bob's measurement. Since $\sum_{v^n} \Lambda_{v^n}^B = I$, from Lemma 3 we have for every $u^n \in \mathcal{T}_\delta^{(n)}(A)$,

$$\sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\left(\frac{1}{\lambda_{u^n}^A} \Lambda_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \tilde{\rho}_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1$$
$$= \left\| \sqrt{\rho_A^{\otimes n}} \left(\frac{1}{\lambda_{u^n}^A} \Lambda_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \tilde{\rho}_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 .$$

This simplifies $\mathbb{E}[J_2]$ as

$$\mathbb{E}[J_2] \leqslant \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \left\| \sqrt{\rho_A^{\otimes n}} \left(\frac{1}{\lambda_{u^n}^A} \Lambda_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \tilde{\rho}_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1$$
$$= \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \left\| (\hat{\rho}_{u^n}^A - \tilde{\rho}_{u^n}^A) \right\|_1$$
$$\leqslant \frac{(1-\varepsilon)}{(1+\eta)} (2\sqrt{\varepsilon_A'} + 2\sqrt{\varepsilon_A''}) = \epsilon_{J_2},$$

where the last inequality is obtained by the repeated usage of the average gentle measurement lemma by setting $\epsilon_{J_2} = \frac{(1-\varepsilon_p)}{(1+\eta)}(2\sqrt{\varepsilon_A'} + 2\sqrt{\varepsilon_A''})$ with $\epsilon_{J_2} \searrow 0$ as $n \to \infty$ and $\varepsilon_A' = \varepsilon_p + 2\sqrt{\varepsilon_p}$ and $\varepsilon_A'' = 2\varepsilon_p + 2\sqrt{\varepsilon_p}$ for $\varepsilon_p \triangleq 1 - \min\left\{ \operatorname{Tr} \Pi_{\rho_A} \hat{\rho}_{u^n}^A, \operatorname{Tr} \Pi_{u^n} \hat{\rho}_{u^n}^A, 1-\varepsilon \right\}$. Since $Q_1 \leqslant J \leqslant J_1 + J_2$, hence J, and consequently Q_1, can be made arbitrarily small for sufficiently large n, if $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$.

F. Proof of Proposition 7

We start by adding and subtracting the following terms in Q_2

$$(i) \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n)$$

$$(ii) \sum_{u^n,v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \times P_{Z|U,V}^n(z^n|u^n,v^n)$$

$$(iii) \sum_{u^n,v^n} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \times P_{Z|U,V}^n(z^n|u^n,v^n).$$

This gives us $Q_2 \leqslant Q_{21} + Q_{22} + Q_{23} + Q_{24}$, where the terms on the right hand side are shown at the bottom of the page. We start by analyzing Q_{21}. Note that Q_{21} is exactly same as Q_1 and hence using the same rate constraints as Q_1, this term can be bounded. Next, consider Q_{22}. Substitution of $\zeta_{v^n}^{(\mu_2)}$ gives

$$Q_{22} = \left\| \sum_{u^n,v^n,z^n} \lambda_{u^n,v^n}^{AB} \hat{\rho}_{u^n,v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n,v^n) |z^n\rangle\langle z^n| - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_2+C_2)}} \sum_{\mu_2,k} \sum_{u^n,v^n,z^n} \mathbb{1}_{\{V^{n,(\mu_2)}(k)=v^n\}} \times \frac{\lambda_{u^n,v^n}^{AB}}{\lambda_{v^n}^B} \hat{\rho}_{u^n,v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n,v^n) |z^n\rangle\langle z^n| \right\|_1,$$

where the equality uses the triangle inequality for block operators. From here on, we use Lemma 8 to bound Q_{22}. For this, let us define $\mathcal{T}_{v^n}$ as

$$\mathcal{T}_{v^n} \triangleq \sum_{u^n,z^n} \frac{\lambda_{u^n,v^n}^{AB}}{\lambda_{v^n}^B} \hat{\rho}_{u^n,v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n,v^n) |z^n\rangle\langle z^n| .$$

Note that $\mathcal{T}_{v^n}$ can be written in tensor product form as $\mathcal{T}_{v^n} = \bigotimes_{i=1}^n \mathcal{T}_{v_i}$. This simplifies Q_{22} as

$$Q_{22} = \left\| \sum_{v^n} \lambda_{v^n}^B \mathcal{T}_{v^n} - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_2+C_2)}} \sum_{\mu_2,k} \mathcal{T}_{V^{n,(\mu_2)}(k)} \right\|_1 .$$

Using Lemma 8, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}(Q_{22}) \leqslant \epsilon$, if

$$\tilde{R}_2 + C_2 > S\left(\sum_{v \in \mathcal{V}} \lambda_v^B \mathcal{T}_v \right) - \sum_{v \in \mathcal{V}} \lambda_v^B S(\mathcal{T}_v) = I(RZ;V)_{\sigma_3}, \tag{76}$$

where σ_3 is defined in the statement of the theorem.

Now, we move on to consider Q_{23}. Taking expectation with respect to the codebook $\mathcal{C}^{(\mu_1,\mu_2)} = (\mathcal{C}_1^{(\mu_1)}, \mathcal{C}_2^{(\mu_2)})$ gives bounds, shown at the top of the next page, where the inequality

$$Q_{21} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\left(\frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \right) \otimes \Lambda_{v^n}^B - \Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$
$$Q_{22} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \Lambda_{u^n}^A \otimes \left(\frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$
$$Q_{23} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \left(\frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) - \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$
$$Q_{24} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B - A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1 .$$

$$\begin{aligned}\mathbb{E}\left[Q_{23}\right] &\leqslant \mathbb{E}_{\mathcal{C}}\left[\sum_{z^n,v^n}\frac{1}{N_2}\sum_{\mu_2=1}^{N_2}\frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right.\right.\\&\qquad\left.\left.-\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1\right]\\&=\mathbb{E}_{\mathcal{C}_1}\left[\sum_{z^n,v^n}\frac{1}{N_2}\sum_{\mu_2=1}^{N_2}\frac{\mathbb{E}_{\mathcal{C}_2}\left[\zeta_{v^n}^{(\mu_2)}\right]}{\lambda_{v^n}^B}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right.\right.\\&\qquad\left.\left.-\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1\right]\\&=\mathbb{E}_{\mathcal{C}_1}\left[\sum_{z^n,v^n}\frac{1}{(1+\eta)}\left\|\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{u^n}^A\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right.\right.\\&\qquad\left.\left.-\sum_{u^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|U,V}^n(z^n|u^n,v^n)\right\|_1\right]\\&=\mathbb{E}\left[\frac{J}{(1+\eta)}\right],\end{aligned}$$

is obtained by using the triangle inequality, and the first equality follows as $\mathcal{C}_1^{(\mu_1)}$ and $\mathcal{C}_2^{(\mu_2)}$ are generated independently. The last equality follows from the definition of J as in (53). Hence, we use the result obtained in bounding $\mathbb{E}[J]$ in the proof of Proposition 6.

Finally, we consider Q_{24}.

$$\begin{aligned}Q_{24}&\leqslant\sum_{u^n,v^n}\sum_{z^n}P_{Z|U,V}^n(z^n|u^n,v^n)\\&\quad\times\left\|\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\mu_1)}\otimes\frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B}\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right.\\&\quad-\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(A_{u^n}^{(\mu_1)}\otimes\zeta_{v^n}^{(\mu_2)}\right.\\&\quad\left.\left.\times\left(\sqrt{\rho_B^{\otimes n}}^{-1}\tilde\rho_{v^n}^B\sqrt{\rho_B^{\otimes n}}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\\&\leqslant\frac{1}{N_2}\sum_{\mu_2}\sum_{u^n,v^n}\zeta_{v^n}^{(\mu_2)}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\frac{1}{\lambda_{v^n}^B}\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right.\\&\quad\left.-\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\left(\sqrt{\rho_B^{\otimes n}}^{-1}\tilde\rho_{v^n}^B\sqrt{\rho_B^{\otimes n}}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1,\end{aligned}$$

where the inequalities above are obtained by substituting in the definition of $B_{v^n}^{(\mu_2)}$ and using multiple triangle inequalities. Taking expectation of Q_{24} with respect to the second codebook generation, we get

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_2}\left[Q_{24}\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right]\\&\leqslant\mathbb{1}_{\{\text{sP-1}\}}\sum_{\substack{u^n\\v^n\in\mathcal{T}_\delta^{(n)}(V)}}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\frac{1}{\lambda_{v^n}^B}\Lambda_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}}\right.\\&\quad\left.-\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\otimes\left(\sqrt{\rho_B^{\otimes n}}^{-1}\tilde\rho_{v^n}^B\sqrt{\rho_B^{\otimes n}}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\\&\overset{(a)}{\leqslant}\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\sqrt{\rho_B^{\otimes n}}\left(\frac{1}{\lambda_{v^n}^B}\Lambda_{v^n}^B\right.\right.\\&\quad\left.\left.-\sqrt{\rho_B^{\otimes n}}^{-1}\tilde\rho_{v^n}^B\sqrt{\rho_B^{\otimes n}}^{-1}\right)\sqrt{\rho_B^{\otimes n}}\right\|_1\\&=\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\hat\rho_{v^n}^B-\tilde\rho_{v^n}^B\right\|_1\\&\overset{(b)}{\leqslant}\frac{(1-\varepsilon')}{(1+\eta)}(2\sqrt{\varepsilon_B'}+2\sqrt{\varepsilon_B''})=\epsilon_{Q_{24}},\end{aligned}\tag{77}$$

where (a) follows by using Lemma 3 and the fact that under the event $\{\mathbb{1}_{\{\text{sP-1}\}}=1\}$ we have $\sum_{u^n}\frac{1}{N_1}\sum_{\mu_1}A_{u^n}^{(\mu_1)}\leqslant I$, and (b) uses the result based on the average gentle measurement lemma by setting $\epsilon_{Q_{24}}=\frac{(1-\varepsilon')}{(1+\eta)}(2\sqrt{\varepsilon_B'}+2\sqrt{\varepsilon_B''})$ with $\epsilon_{Q_{24}}\searrow 0$ as $n\to\infty$ and $\varepsilon_B'=\varepsilon_p+2\sqrt{\varepsilon_p}$ and $\varepsilon_B''=2\varepsilon_p+2\sqrt{\varepsilon_p}$, for $\varepsilon_p\triangleq 1-\min\left\{\operatorname{Tr}\Pi_{\rho_B}\hat\rho_{v^n}^B,\operatorname{Tr}\Pi_{v^n}\hat\rho_{v^n}^B,1-\varepsilon'\right\}$. This completes the proof for Q_{24} and hence for all the terms corresponding to Q_2.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their comments that greatly helped in preparing a better manuscript. They also thank Mark Wilde for his valuable inputs on techniques needed to prove Theorem 8 and for referring them to the additional work performed in [10] and [47]. They are also grateful to Arun Padakandla for his inputs on the classical analogue of the current work [48], which was very helpful in developing the proof techniques here.

REFERENCES

[1] A. Winter, "'Extrinsic,' and 'intrinsic' data in quantum measurements: Asymptotic convex decomposition of positive operator valued measures," *Commun. Math. Phys.*, vol. 244, no. 1, pp. 157–185, 2004.

[2] M. M. Wilde, P. M. Hayden, F. Buscemi, and M.-H. Hsieh, "The information-theoretic costs of simulating quantum measurements," *J. Phys. A, Math. Theor.*, vol. 45, no. 45, Nov. 2012, Art. no. 453001.

[3] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2926–2959, May 2014.
[4] H. J. Groenewold, "A problem of information gain by quantal measurements," *Int. J. Theor. Phys.*, vol. 4, no. 5, pp. 327–338, Sep. 1971.
[5] G. Lindblad, "An entropy inequality for quantum measurements," *Commun. Math. Phys.*, vol. 28, no. 3, pp. 245–249, Sep. 1972.
[6] M. Ozawa, "On information gain by quantum measurements of continuous observables," *J. Math. Phys.*, vol. 27, no. 3, pp. 759–763, Mar. 1986.
[7] F. Buscemi, M. Hayashi, and M. Horodecki, "Global information balance in quantum measurements," *Phys. Rev. Lett.*, vol. 100, no. 21, 2008, Art. no. 210504.
[8] S. Luo, "Information conservation and entropy change in quantum measurements," *Phys. Rev. A, Gen. Phys.*, vol. 82, no. 5, 2010, Art. no. 052103.
[9] M. E. Shirokov, "Entropy reduction of quantum measurements," *J. Math. Phys.*, vol. 52, no. 5, May 2011, Art. no. 052202.
[10] M. Berta, J. M. Renes, and M. M. Wilde, "Identifying the information gain of a quantum measurement," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7987–8006, Dec. 2014.
[11] M. Horodecki, J. Oppenheim, and A. Winter, "Partial quantum information," *Nature*, vol. 436, pp. 673–676, Aug. 2005.
[12] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," *Commun. Math. Phys.*, vol. 269, no. 1, pp. 107–136, 2007.
[13] M. Christandl, R. König, and R. Renner, "Postselection technique for quantum channels with applications to quantum cryptography," *Phys. Rev. Lett.*, vol. 102, no. 2, 2009, Art. no. 020504.
[14] A. Anshu, R. Jain, and N. A. Warsi, "Convex-split and hypothesis testing approach to one-shot quantum measurement compression and randomness extraction," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5905–5924, Sep. 2019.
[15] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum communication using coherent rejection sampling," *Phys. Rev. Lett.*, vol. 119, Sep. 2017, Art. no. 120506.
[16] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum message compression with applications," 2014, *arXiv:1410.3031*.
[17] A. Anshu, R. Jain, and N. A. Warsi, "Building blocks for communication over noisy quantum networks," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1287–1306, Feb. 2019.
[18] J. M. Renes and R. Renner, "One-shot classical data compression with quantum side information and the distillation of common randomness or secret keys," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1985–1991, Mar. 2012.
[19] M. Tomamichel, *Quantum Information Processing with Finite Resources—Mathematical Foundations* (Springer Briefs in Mathematical Physics), vol. 5. Heidelberg, Germany: Springer, 2015.
[20] S. Khatri and M. M. Wilde, "Principles of quantum communication theory: A modern approach," 2020, *arXiv:2011.04672*.
[21] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, Oct. 2002.
[22] M. Berta, M. Christandl, and R. Renner, "The quantum reverse Shannon theorem based on one-shot information theory," *Commun. Math. Phys.*, vol. 306, p. 579, Sep. 2011.
[23] M. Horodecki *et al.*, "Local information as a resource in distributed quantum systems," *Phys. Rev. Lett.*, vol. 90, no. 10, Mar. 2003, Art. no. 100402.
[24] M. Horodecki *et al.*, "Local versus nonlocal information in quantum-information theory: Formalism and phenomena," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 6, Jun. 2005, Art. no. 062307.
[25] I. Devetak, "Distillation of local purity from quantum states," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 6, Jun. 2005, Art. no. 062303.
[26] H. Krovi and I. Devetak, "Local purity distillation with bounded classical communication," *Phys. Rev. A, Gen. Phys.*, vol. 76, no. 1, Jul. 2007, Art. no. 012321.
[27] J. Devetak, A. W. Harrow, and A. J. Winter, "A resource framework for quantum Shannon theory," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4587–4618, Oct. 2008.
[28] N. Datta, M.-H. Hsieh, M. M. Wilde, and A. Winter, "Quantum-to-classical rate distortion coding," *J. Math. Phys.*, vol. 54, no. 4, 2013, Art. no. 042201.
[29] N. Datta, M.-H. Hsieh, and M. M. Wilde, "Quantum rate distortion, reverse Shannon theorems, and source-channel separation," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 615–630, Jan. 2013.
[30] I. Devetak and A. Winter, "Classical data compression with quantum side information," *Phys. Rev. A, Gen. Phys.*, vol. 68, no. 4, Oct. 2003, Art. no. 042301.
[31] T. Berger, "Multiterminal source coding," in *The Information. Theory Approach to Communications*, G. Longo, Ed. New York, NY, USA: Springer-Verlag, 1977.
[32] A. Yimsiriwattana and S. J. Lomonaco, Jr., "Distributed quantum computing: A distributed shor algorithm," *Proc. SPIE*, vol. 5436, pp. 360–372, Aug. 2004.
[33] R. Beals *et al.*, "Efficient distributed quantum computing," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 469, no. 2153, 2013, Art. no. 20120686.
[34] R. Van Meter and S. J. Devitt, "The path to scalable distributed quantum computing," *Computer*, vol. 49, no. 9, pp. 31–42, Sep. 2016.
[35] V. S. Denchev and G. Pandurangan, "Distributed quantum computing: A new frontier in distributed systems or science fiction?" *ACM SIGACT News*, vol. 39, no. 3, pp. 77–95, Sep. 2008.
[36] A. Anshu, R. Jain, and N. A. Warsi, "A generalized quantum Slepian–Wolf," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1436–1453, Mar. 2018.
[37] M. M. Wilde, private communication, Aug. 2019.
[38] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction*, vol. 16. Berlin, Germany: Walter de Gruyter, 2012.
[39] M. M. Wilde, "From classical to quantum Shannon theory," 2011, *arXiv:1106.1445*.
[40] F. Shirani and S. S. Pradhan, "Finite block-length gains in distributed source coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 1702–1706.
[41] A. B. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.
[42] D. S. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
[43] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[44] G. M. Ziegler, *Lectures on Polytopes* (Graduate Texts in Mathematics), vol. 152. New York, NY, USA: Springer-Verlag, 1995.
[45] J. B. Conway, *A Course in Functional Analysis*. New York, NY, USA: Springer, 1985.
[46] P. A. Meyer, *Quantum Probability for Probabilists* (Lecture Notes in Mathematics), vol. 1538, 2nd ed. Berlin, Germany: Springer-Verlag, 1995.
[47] H. Martens and W. M. de Muynck, "Nonideal quantum measurements," *Found. Phys.*, vol. 20, no. 3, pp. 255–281, Mar. 1990.
[48] T. A. Atif, A. Padakandla, and S. S. Pradhan, "Source coding for synthesizing correlated randomness," 2020, *arXiv:2004.03651*.

Touheed Anwar Atif (Graduate Student Member, IEEE) received the B.Tech. degree from the Department of Electrical Engineering, Indian Institute of Technology Hyderabad, in 2017, and the M.Sc. degree in electrical engineering from the University of Michigan, Ann Arbor, in 2019, where he is currently pursuing the Ph.D. degree in electrical engineering and computer science. His research interests include quantum information theory and quantum algorithms.

Mohsen Heidari received the B.Sc. and M.Sc. degrees in electrical engineering from the Sharif University of Technology in 2011 and 2013, respectively, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, in 2017 and 2019, respectively. He is currently a Visiting Assistant Professor with the Department of Computer Science, Purdue University. His research interests lie in the intersection of information theory, quantum information theory, and learning theory.

S. Sandeep Pradhan (Fellow, IEEE) received the M.E. degree from the Indian Institute of Science in 1996 and the Ph.D. degree from the University of California at Berkeley in 2001. From 2002 to 2008, he was an Assistant Professor with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, where he was also an Associate Professor, from 2008 to 2015. He is currently a Professor with the Department of Electrical Engineering and Computer Science, University of Michigan. His research interests include network information theory, coding theory, and quantum information theory. He was an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY in the area of Shannon theory from 2014 to 2016.